EE 61-6

INTERIM REPORT

PURDUE UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

Present State of the Art in the Specification of Nonlinear Conrrol Systems

J. E. Gibson, Principal Investigator
E. S. McVey, Assistant Principal Investigator
C. D. Leedham
Z. V. Rekasius
D. G. Schultz
R. Sridhar

Control and Information Systems Laboratory

May, 1961
Lafayette, Indiana

RESEARCH PROJECT PRF 2030
CONTRACT NO AF 29(600)-1933
FOR
UNITED STATES AIR FORCE
AIR FORCE MISSILE DEVELOPMENT CENTER
HOLLOMAN AIR FORCE BASE
NEW MEXICO

Interim Report 2

PRESENT STATE OF THE ART IN THE SPECIFICATION OF NONLINEAR CONTROL SYSTEMS

J. E. Gibson, Principal Investigator

E. S. McVey, Assistant Principal Investigator

C. D. Leedham

Z. V. Rekasius

D. G. Schultz

R. Sridhar

SCHOOL OF ELECTRICAL ENGINEERING

PURDUE UNIVERSITY

Lafayette, Indiana

MAY, 1961

UNITED STATES AIR FORCE

AIR FORCE MISSILE DEVELOPMENT CENTER

HOLLOMAN AIR FORCE BASE

New Mexico

PREFACE

This report was prepared by Purdue University, School of Electrical Engineering, Prof. J. E. Gibson acting as Principal Investigator, under USAF Contract No. AF 29(600)-1933. This contract is administered under the direction of the Guidance and Control Division, Air Force Missile Development Center, Holloman Air Force Base, New Mexico by Mr. J. H. Gengelbach, the initiator of the study.

FOREWORD

This is the fifth and last report to be completed under the Air Force Project Number AF 29(600)-1933. The task specified under the above contract is the specification of linear and nonlinear control systems. Toward the achievement of this purpose, the first three reports deal with linear control systems, while the last two concentrate on nonlinear systems.

Interim Report #1, titled "Specification and Data Presentation in Linear Control Systems" was issued in July of 1959. This report was circulated through the control industry and the universities, and a number of the leading industrial concerns in the country were visited in connection with the contents of this report. As a result of this feedback, the basic material of this interim report was expanded and published in two final reports, namely, Final Report, Volume I, "Specification and Data Presentation in Linear Control Systems," October 1960, and Final Report, Volume II, "Specification and Data Presentation in Linear Control Systems, - Part Two," May, 1961. These volumes also carry the Air Force Designation AFMDC-TR-61-5, Parts One and Two. The first of these final reports deals with the specification of continuous systems which can be described by linear differential equations with constant coefficients. The second considers Sampled Data Systems, Linear Time Variable Parameter Systems, and Performance Indicies.

Final Report, Volume III is a tutorial report titled "Stability of Nonlinear Control Systems by the Second Method of Liapunov," and dated May, 1961, (AFMDC-TR-61-6). This report was written to acquaint the interested reader with a technique, common in the USSR, that will serve as a tool in the future nonlinear work, and not as a direct attack on the nonlinear control specification problem.

The present report is an interim report which reviews the status of the nonlinear control art, and specifically the area of nonlinear control system specification. While the complexity of this problem is at least an order of magnitude greater than in the linear case, it is felt that the ideas presented here form the foundation from which a more detailed and explicit attack on the general nonlinear specification problem may be built.

ABSTRACT

This is an interim report on the specification of nonlinear automatic control systems. ~~It~~ This report is concerned primarily with assessing the state of the art of nonlinear control as a prelude to the solution of the actual specification problem.

As an introduction, the classical methods of nonlinear analysis are discussed, and the reasons for the inadequacy of these techniques for automatic control systems are explained. The two generally known methods of analyzing the stability of autonomous nonlinear control systems, namely phase plane analysis and the describing function, are discussed and a summary of the capabilities and limitations of ~~Liapunov's Second~~ Method is presented. The concept of the state variable and the state space is introduced in some detail, as it is expected that this will be the medium through which the stability and response of the majority of nonlinear systems will be handled. The stability of the nonautonomous system and is also discussed from the point of view of signal stabilization and the dual input describing function.

It is pointed out that in addition to the stability of a nonlinear system, its response to a given input is of particular interest. Chapters 3 and 4 are devoted to the response of autonomous and nonautonomous systems. As a criterion for specification, the time optimum system is stressed, and distinction is made between the solution of the time optimum problem as a performance index and the synthesis of the optimum switching boundaries. The phase plane is discussed for forced systems and the work of Wiener is mentioned in connection with the response of nonlinear systems to random inputs.

TABLE OF CONTENTS

TABLE OF CONTENTS - CONTINUED

LIST OF ILLUSTRATIONS

CHAPTER I

INTRODUCTION TO NONLINEAR CONTROL SYSTEMS

1.1 Introduction

All physical systems are nonlinear, although in many systems this nonlinear effect is so slight that satisfactory results are obtained with linear models. Many physical systems are nonlinear simply due to the lack of component perfection. However, a significant number of systems are nonlear through conscious design. Many times a nonlinear system will be lighter, cheaper, more reliable, easier to fabricate and have better performance than an equivalent linear system. Thus it is of great importance to the Air Force that nonlinear control systems be properly specified.

This is an interim report on the specification of nonlinear automatic control systems. It has three objectives:

a) to show why classical nonlinear mechanics has not provided the tools needed by the automatic control engineer thus far.

b) to assessthe present state of the nonlinear automatic control art.

c) to point out the directions future work will take.

1.2 Classical Nonlinear Mechanics and Nonlinear Analysis as Applied to Automatic Control

Classical nonlinear mechanics has generally been used for the analysis of nonlinear problems. For some few problems it has been possible to find closed form solutions in terms of the simpler functions. Generally, however, this attack fails. A number of books have been written on special nonlinear differential equations and it would not be difficult to fill report after report with such considerations. This is not necessary, however,

nor would it even be proper, since seldom, if ever, will it be possible to arrange even a moderately complex control system into a form which would make use of available solutions. This is not to say, of course, that a background of such techniques will not prove valuable to a designer. In fact it is obvious that in a difficult field such as nonlinear automatic control it is desirable to have as much training as possible.

It has been long realized, of course, that closed form solutions to nonlinear differential equations are difficult to obtain and exist for only a few special classes. For 100 years or more analysts have been concerned with the approximate solution of nonlinear differential equations. Such series approximation techniques as perturbation and reversion are well known. Other methods such as variation of parameters and harmonic balance are also widely used. The mathematical justification of these methods generally requires that the nonlinear variation be small and/or slow and/or smooth. Sometimes the engineer is faced with nonlinear control systems in which none of these restrictions are valid and simulation is the only practical solution. It is apparent that classical exact solutions are of little value.

A number of excellent texts are available that will introduce the engineer to nonlinear analysis. The recent book by Cunningham [1] is notable for the clarity of presentation and the numerous worked examples. Other well known books are those by Stoker [2], Minorsky [3] and Andronow and Chaikin [4]. Somewhat more intense mathematically are the books by Lefschetz [5] and Bellman [6].

1.3 Approximate Methods for Nonlinear Control

Modern approximate techniques of nonlinear system analysis are direct

outgrowths of classical analysis, and one could probably relate them directly to Poincare' and Liapunov, if there was any point in so doing. This discussion will be avoided by pointing out that the newness lies in the emphasis and phrasing of the problem and the prominence of geometric and graphical interpretation, but not in techniques of analysis.

Chapter 2 of this report considers the more important of these techniques in detail, so they need not be discussed here.

The state of the art in the analysis and synthesis of nonlinear control systems is unsatisfactory, especially in its lack of generality. It is almost impossible to rely on a single analysis to illustrate all of the possible phenomena that can occur in a single system. For example, it takes a different analysis to demonstrate jump phenomena than it does to show subharmonic oscillation or frequency entrainment for the same system. It does not appear that this condition will change in the near future, because approximate analysis is not completely reliable, and some other method must be used to supplement the analysis of a nonlinear control system. This other approach, widely used now, is computer simulation which is discussed in the next section.

1.4 Computer Simulation

The major emphasis in this report is on analysis, because it is desired to obtain an understanding of systems in general to facilitate the evaluation and specification problem. However, it is recognized that engineers use computer simulation for nonlinear system analysis more than they use mathematical methods. There are several reasons for this.

1. Mathematical methods are not available or are not tractable for the determination of system response. It is usually less ex-

pensive to obtain a transient response on a computer than by analysis.

2. In a nonlinear system complete knowledge of any particular response does not necessarily imply knowledge of any other response. Thus, it may be necessary to obtain thousands of responses to establish confidence in a design. This makes all but the simplest calculations uneconomical.
3. Actual systems are often much more difficult to analyze than simple text book examples. It may be necessary to include some actual pieces of hardware in the simulation if they can not be described adequately. Analysis, of course, is not this flexible.
4. Engineers who do design work may not be aware of the mathematical tools available for design and evaluation.
5. Design engineers are usually more interested in a specific system than general trends which are available from mathematical analysis. Hence, a thorough simulation is often adequate for their purposes. For example, the parameters of a simulated system can be varied and the resulting response observed for system synthesis.

The present inadequacy of nonlinear analysis should not lead one to abandon all attempts at analysis and to a complete reliance on computer simulation; a combination of simulation and analysis seems more nearly optimum than simulation alone. Some analysis, even with incomplete or inexact models, will yield insight not always available from simulation. Major advances in theory, and hence hardware, will be delayed if attention is not given to the mathematical treatment of systems.

Thus the tentative recommendation of the Purdue group will no doubt in-

volve a parallel use of computer simulation and modern approximate analytical techniques for the specification of nonlinear control systems.

1.5 Future Trends

In each of the following chapters an assessment of the importance upon future developments of the techniques discussed is given. In fact, because of the present incompleteness of existing techniques, a great deal of space in this report seems to be given over to damning the status quo. This may be interpreted as an undesirable situation and to provide cause for discouragement. The Control and Information Systems Laboratory, on the other hand, feels that specifically pointing out the deficiencies means that we have at least progressed to the point where we recognize the problem. This could not have been said of most automatic control engineers as late as 4 or 5 years ago.

The reader of this report, especially if he has been concerned with automatic control systems for a decade or more, will recognize an almost revolutionary change in techniques and emphasis compared with what might be called, "Classical Automatic Control". This is the collection of techniques available in almost all of the texts in English. The "New Automatic Control" is more advanced mathematically and calls upon the digital computer as an on line element more-and-more frequently. It works frequently in a non-physical state space and attempts to find the theoretical limits of performance based upon ultimate physical limitations on the system, such as finite energy or torque or velocity, but without consideration of the detailed construction of any particular configuration. In other words, the optimum problem becomes important. The ultimate time optimum systems are studied and the self optimizing or adaptive problem is of concern.

The "New Automatic Control" is, as yet, essentially an academic discipline. The reader will see that few if any practical systems have benefited as yet from this approach. However, only 5 to 10 years after the classical automatic control matured in the early 1940's, it became an essential part of engineering system design. It seems entirely possible that the 1960's will witness a similar impact on industrial and aerospace system design due to the "New Automatic Control".

When reading some of the mathematical work contained in the report, the reader should keep in mind that a mathematical treatment of a problem is usually the starting point for engineering effort, rather than a practical problem solution. For example, while the formal solution is desired for the general, time varying, optimum, switched system problem, it must be realized that practical, general problems are either not mathematically tractable or are trivial. In addition, it should be pointed out that practical aspects of the problem such as end point switching, instrument imperfections, etc. have not been included in the general formulation. This single example serves the purpose of illustrating the obvious -- much research remains to be done in the nonlinear area of control systems.

CHAPTER II

STABILITY OF AUTONOMOUS SYSTEMS

2.1 Introduction

The word "stability" is frequently interpreted by engineers as that property of a system which yields a bounded response to any bounded input or load disturbance. While such interpretation is correct in linear, stationary systems, it may easily lead one to erroneous conclusions in the case of nonlinear systems. In nonlinear stationary systems the "boundedness" of response to bounded inputs no longer guarantees that the unforced system response will return to the equilibrium state asymptotically in time. Neither is the converse true (i.e., asymptotic stability does not always imply total stability or stability in the presence of bounded inputs and/or load disturbances).

Additional complications arise due to the fact that in nonlinear systems stability of an equilibrium state is no longer a global concept but only a local system property (i.e., a nonlinear system may be stable for sufficiently small initial disturbances and become unstable because of a sufficiently large disturbance, and vice versa). Furthermore, it is conceivable that a nonlinear system may be stable for certain bounded inputs and become unstable for other bounded inputs. Hence, in the analysis, synthesis, and specification of nonlinear automatic control systems, total stability (i.e., stability in the presence of any bounded input or disturbance) is the ultimate (although not always necessary) goal. Nevertheless there are several important reasons why the stability of autonomous (unforced, stationary) control systems is of considerable importance:

1) It is important to know the behaviour of the system in the absence of inputs and load disturbances.
2) In the presence of constant inputs and/or constant load disturbances a nonlinear control system can still be described by a set of autonomous differential equations.
3) Stability of the equilibrium state (i.e., stability in Liapunov sense) or boundedness of unforced stationary system response (i.e., stability in Lagrange sense (La Salle [7]) implies boundedness of the response to bounded inputs or total stability in most (if not all) physical systems.

This chapter is devoted to a discussion of the more general or more promising methods of stability analysis of nonlinear autonomous systems.

2.2 The Describing Function Method of Analysis

The describing function (D.F.) method of analysis is appealing from a practical point of view because it is an attempt to linearize a certain class of nonlinear systems and then apply the methods of linear system stability analysis. Engineers are accustomed to making simplifying assumptions and using linearized models for the analysis and synthesis of nonlinear systems. The D.F. is based on the method of harmonic balance (Kryloff [8]), (Cunningham [1]). Several papers (Goldfarb [9]), (Kochenburger [10]), (Tustin [11]), (Oppeit [12]) have advanced this idea. The most common describing function, or the so-called equivalent gain, is defined as the complex ratio of the amplitude of the fundamental component of the output of a nonlinearity to the amplitude of the input to the nonlinearity when the input is sinusoidal. Restrictions such as low pass filtering must be met by the system for the analysis to be valid. A detailed discussion of the method is unnecessary here

since the D.F. is one of the most well known methods available for the analysis of autonomous nonlinear systems.

The D.F. method can be used to determine the stability of an autonomous system and provides a designer with the information necessary to synthesize stabilizing networks. If an autonomous system has a stable limit cycle, the approximate amplitude and frequency of the first harmonic term of the oscillation can be predicted. It is possible to obtain higher harmonic correction terms (Johnson [13]) which improve the accuracy of the method. The work required to calculate the correction terms is generally not justified because these terms are relatively small in systems which possess adequate low-pass filter characteristics. Their main utility is the confidence established in the validity of the D.F. if these correction terms are relatively small. Gille, et al. ([14], p 43) point out that cases are unusual where the error introduced in neglecting the higher harmonic terms exceeds 10 per cent, and that the accuracy of limit cycle frequency obtained from the D.F. is usually better than 5 percent.

Levinson [15] and other investigators (Hill [16]) have used the describing function to predict the closed loop frequency response of stationary nonlinear systems. By writing a quasi-linear error transfer function and solving (usually by means of a computer or graphical techniques) for a value of error which satisfies this quasi-linear transfer function, error is determined. Knowing the error, the response may then be found. Multiple roots of the solutions of the above transfer function yield information about the jump phenomena of the system. This is a laborious process and different results are obtained for different amplitudes of the input, since the system is nonlinear. This method is valid only for nonlinear systems

that are totally stable. If this point is not recognized, it is possible to obtain erroneous results. Further discussions regarding frequency response of nonlinear systems are given in the section on Dual Input Describing functions in the next chapter.

The D.F. method continues to be a vehicle for nonlinear research as well as design. Perhaps the major disadvantage of the method is that it is limited to frequency analysis. Of course, other methods share this deficiency also.

Tsypkin [17] has presented a method equivalent to the D.F. method for the exact analysis of unforced on-off (relay) systems. This method retains all harmonics generated by the nonlinear element. When harmonics are neglected and only the fundamental component of the output of the nonlinear element is used, this method reduces to the conventional describing function method of analysis. The use of this method is not warranted in systems which possess sufficient high-frequency attenuation such that the approximate describing function method of analysis is adequate. Furthermore, the method of Tsypkin is practical only with very simple nonlinearities, such as a relay, and cannot be applied to more general types of nonlinear systems.

2.3 Phase Plane Analysis

The phase plane method of analysis is applicable directly to only second order nonlinear autonomous systems. This method consists of investigating the behavior of the trajectories of system response in the plane of some system variable and its first time derivative. A detailed discussion of the phase plane method of analysis can be found in many textbooks on nonlinear analysis [1].

A generalization of the phase plane analysis is the analysis in the phase space, i.e., in the space of a variable of the system and its n-1 time derivatives where n is the order of the system. Unfortunately, the amount of labor involved in constructing the phase trajectories in systems of higher than second order is prohibitive [18]. Hence the practical use of the phase space (phase plane) method of stability analysis is limited to only the second order autonomous nonlinear systems.

2.4 The Concept of State Space

Before proceeding with the analysis and synthesis of a control system, one has first to find a mathematical description of such a system. In stationary linear systems this is usually accomplished by first expressing the interrelationships between various variables of the system in terms of linear differential equations with constant coefficients. Then these differential equations are changed (by means of the Laplace transform or other integral transforms) into transfer functions and combined to yield an overall transfer function.

In nonlinear systems the Laplace transformation is no longer applicable, and thus the mathematical description of the system must be retained in the form of differential equations. The most convenient form for many purposes is a description of the system by means of n first order differential equations. This can always be done in a straight-forward manner by properly identifying the variables appearing in the system. The number of independent first order differential equations is equal to the order of the system (i.e. to the order of a single differential equation describing the system). The set of n independent first order equations completely describes the state of the system at any time t. Hence a set of n linearly independent

variables will be referred to as a set of state variables and the Euclidean space of these state variables as the state space. One may note that an infinite number of state variable sets may be chosen to represent the same system. Probably the simplest set of state space variables is the set of phase space variables (Sec. 2.3).

To assist in the design and analysis of nonlinear systems a standard form for the differential equations and the system block diagram (if applicable) is used in terms of variables that are not necessarily those of the physical system. The term "canonic form" is used frequently and interchangeably with the term "standard form". It implies one of several of the simplest and most significant forms to which general equations may be brought without loss of generality. The form is mathematically convenient and the advantages of such a form out-weigh the advantages of retaining the system physical variables. It is often convenient, in fact, to write the equations of any system, linear or nonlinear, of high or of low order, in such a canonical form.

The principal characteristic associated with systems in canonical form is that the different variables are "separated", i.e. each of the n first order differential equations contains only one variable, or if this is not possible, some may contain two variables.

A particular form of the system variables may be chosen, therefore, so that the system equations in terms of these variables will reduce to the standard or canonical form. The new variables, $(y_1, \ldots . y_n)$ associated with the canonical form of the system equations, are related by a linear transformation to the system physical variables $(x_1, x_2, \ldots x_n)$ such that:

$$\left.\begin{aligned} x_1 &= P_{11}y_1 + P_{12}y_2 \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot P_{1n}y_n \\ &\vdots \\ x_n &= P_{n1}y_1 + P_{n2}y_2 \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot P_{nn}y_n \end{aligned}\right\} \tag{2.1}$$

or in matrix notation (Pipes [19], Chapter 4)

$$\{x\} = [P]\{y\} \tag{2.2}$$

or

$$\{y\} = [P]^{-1}\{x\} \tag{2.3}$$

where

$[P]$ = a square nxn matrix with elements P_{ij}

$[P]^{-1}$ = inverse of $[P]$

$\{y\}$ = a nx1 matrix with elements y_i

$\{x\}$ = a nx1 matrix with elements x_i

The theory of linear transformations indicates that the basic properties of the system (e.g. the characteristic roots or eigenvalues of the system linear portion) are identical in either set of variables [1], p. 89).

In the language of a positional control system, one new variable could be defined in the form:

$$y_1 = Px + Qv + Ra \tag{2.4}$$

where x = position

v = velocity

a = acceleration

P,Q,R = constants

This example indicates that the physical meaning of the new variables usually is obscure. The mathematical simplification that results is, however, of considerable importance.

With the physical meaning of the new variables obscure, one can, with very little further effort, consider them to be measured in Euclidean n-space

along a set of n mutually perpendicular axes. We have, therefore, the vector:

$$\{y\} = \begin{Bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{Bmatrix} = y_1 a_{y_1} + y_2 a_{y_2} + \cdots + y_n a_{y_n} \tag{2.5}$$

where the a_{y_n} are unit vectors defining the axes in n-space. This vector in n-space describes the state of the system completely.

There are an infinite number of square matrices $[P]$ that will perform a linear transformation on the physical variables, $x_1 \ldots x_n$. The choice of $[P]$ is critical, therefore, in that it defines the canonical form in which the system equations are written.

The procedure that can be followed to select the matrix $[P]$ will be described by using a particular example. Consider the closed-loop system with separable nonlinearity as shown in Figure 2.1. It is assumed that the differential equations for the actual system have been written and expressed in the form shown in this figure. From this form the following relations can be written:

$$E_1(s) = R_1(s) - X_1(s) \tag{2.6}$$

$$\frac{X_1(s)}{U(s)} = \frac{1}{s(s+1)(s+2)} \tag{2.7}$$

which when combined and transformed to the time domain give:

$$\frac{d^3e_1}{dt^3} + \frac{3d^2e_1}{dt^2} + \frac{2de_1}{dt} = (-1)u + \frac{d^3r}{dt^3} + \frac{3d^2r}{dt^2} + \frac{2dr}{dt} \tag{2.8}$$

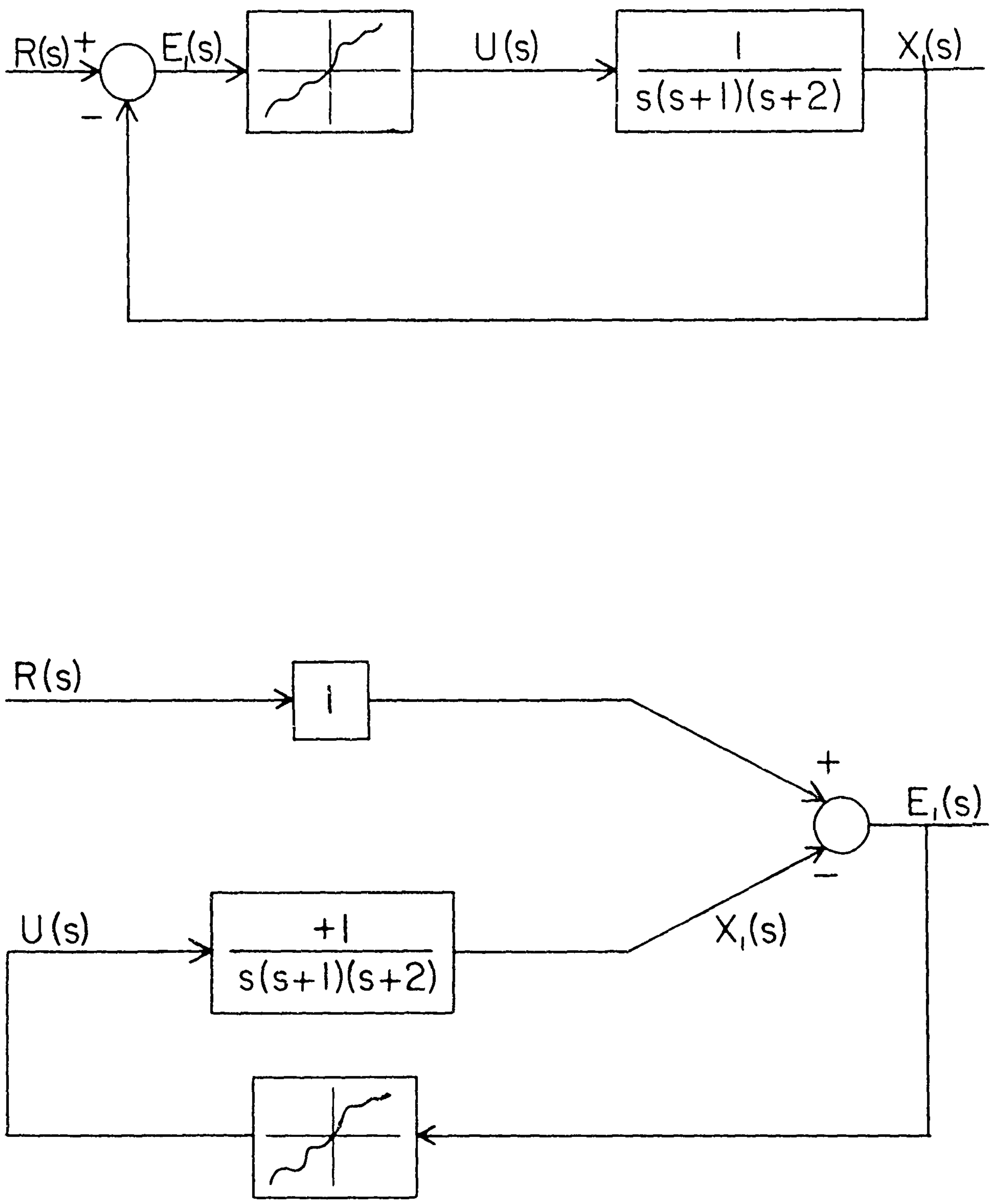

Figure 2.1

The Conventional Block Diagram of a Closed Loop

System with a Separable Nonlinearity (top)

and Equivalent Block Diagram (bottom)

For a given input r(t), the quantity $\frac{d^3r}{dt^3} + \frac{3d^2r}{dt^2} + \frac{2dr}{dt}$ is known and will be abbreviated f(t).

New variables are now introduced,

$$e_2 = \frac{de_1}{dt} \quad \text{and } e_3 = \frac{de_2}{dt} \qquad (2.9)$$

so that equation 2.8 can be re-written:

$$\frac{de_1}{dt} = (0)\ e_1 + (1)e_2 + (0)e_3 + (0)u + (0)f$$

$$\frac{de_2}{dt} = (0)\ e_1 + (0)e_2 + (1)e_3 + (0)u + (0)f \qquad (2.10)$$

$$\frac{de_3}{dt} = (0)\ e_1 + (-2)e_2 + (-3)e_3 + (-1)u + (1)f$$

or

$$\begin{Bmatrix} \dot{e}_1 \\ \dot{e}_2 \\ \dot{e}_3 \end{Bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -3 \end{bmatrix} \begin{Bmatrix} e_1 \\ e_2 \\ e_3 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ -1 \end{Bmatrix} u + \begin{Bmatrix} 0 \\ 0 \\ f \end{Bmatrix} \qquad (2.11)$$

which in matrix notation becomes:

$$\{\dot{e}\} = [A]\{e\} + [B]\{u\} + \{f\} \qquad (2.12)$$

Any systems which are linear in the sense that the elements being controlled are linear and where the steering function, u(t), enters linearly as a function of time can be reduced to a similar form. In this example the $[A]$ matrix, the system matrix, has elements that are constants because the linear portion of the system had constant coefficents. If the linear portion of the system had been time varying the matrix elements would have been time varying. In general the $[B]$ matrix would be nxr and the $\{u\}$ vector of dimension r; in

this example r = 1.

It is emphasized that while any system will reduce to the form of equation (2.12) the details of the equation are not unique for a given system. If Figure 2.1 is rearranged as in Figure 2.2, equation (2.11) becomes:

$$\begin{Bmatrix} \dot{e}'_1 \\ \dot{e}'_2 \\ \dot{e}'_3 \end{Bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -2 \end{bmatrix} \begin{Bmatrix} e'_1 \\ e'_2 \\ e'_3 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix} u + \begin{Bmatrix} g \\ 0 \\ 0 \end{Bmatrix} \tag{2.13}$$

where $g = \frac{dr}{dt}$. This equation obviously has the same form as equation (2.11) but differs in detail.

Thus far the variables used are close to the physical variables, though they may not be available directly in the system. Should the system have zero input the equations can be written directly in terms of the output and its derivatives, $x_1 \ldots x_n$. Returning to Figure 1 and setting $r(t) = 0$, the following equation can be derived to replace equation (2.11).

$$\begin{Bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{Bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -3 \end{bmatrix} + \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ +1 \end{Bmatrix} u \tag{2.14}$$

or

$$\{\dot{x}\} = [A]\{x\} + [B_1]\{u\} \tag{2.15}$$

The details of equation (2.14) are, of course, not unique either.

The new variables $\{y\}$ associated with the state space to be used and related to the existing variables $\{x\}$ or $\{e\}$ by equations (2.7) and (2.8)

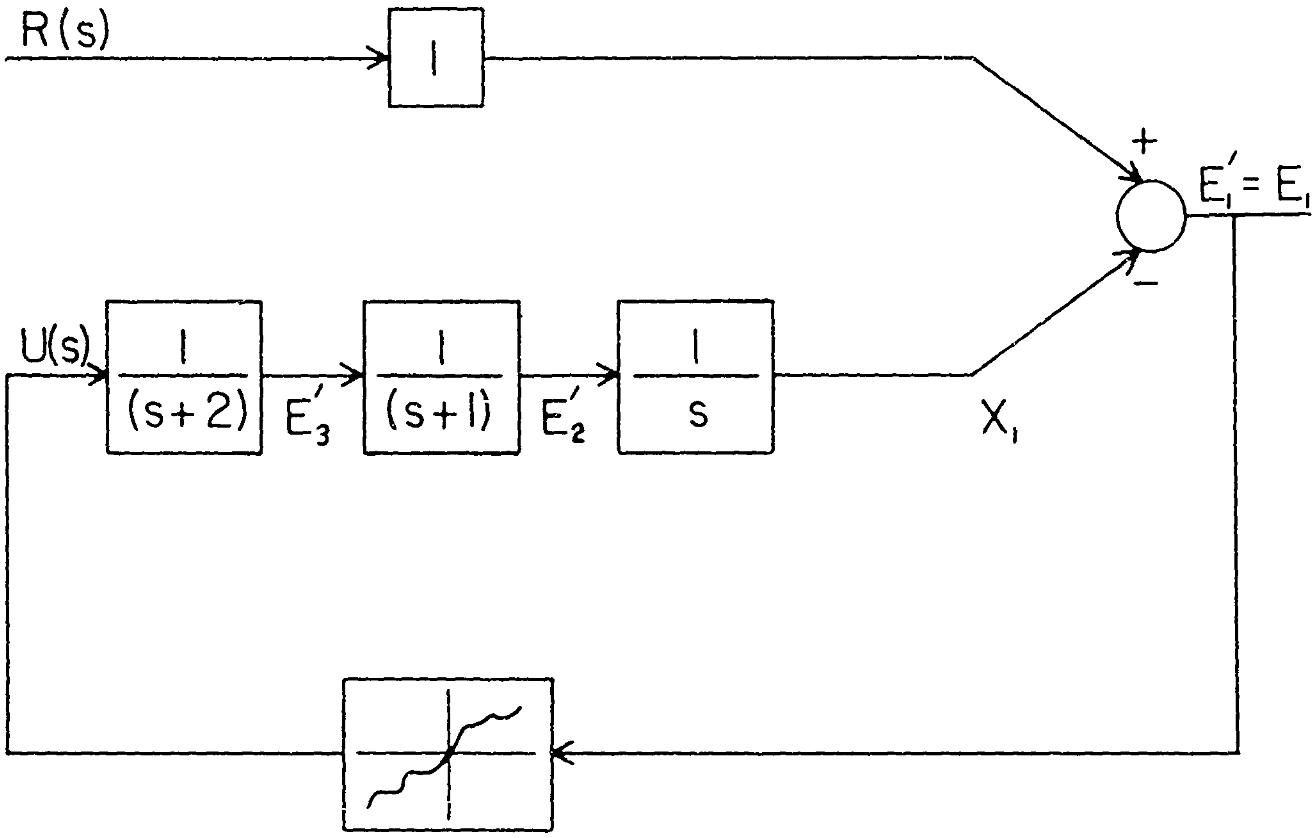

Figure 2.2

The System of Figure 2.1 Redrawn

must now be found. The form of the matrix $[P]$ must be determined on the basis that the equations will transform into a form that is mathematically convenient. There are a number of techniques available that will yield the elements of this matrix. Other methods yield, directly, the system impulse response matrix $[H]$ which will be used and defined later in this section. The methods are, for example: The classical method of separation of variables ([19], Chapter 4, Section 20): The general solution by Lagrange's method of variation of parameters which yields the $[H]$ matrix directly (La Salle [20], Section 2), (Bellman [21], Chapter 10, Section 12): The method reported by Kalman [22]: Lur'e's canonical form and the psuedo canonical form ([23] Chapter 2): Solution in terms of the Jordan canonical form (Kaplan [24], p. 289): A summary and variations on several methods by Kurzweil [25].

The example system chosen, described by equations (2.11) and (2.14), is characterized by the fact that the linear portion has real, distinct poles, i.e. the system matrix $[A]$ has real, distinct eigenvalues given by the solution of the equation $\left|[A] - \lambda[I]\right| = 0$ ([1] p. 88). The method to use in the determination of the matrix $[P]$ depends, as is described in the above mentioned references, on the form of the system differential equations. In this case, the example of equations (2.11) and (2.14), the classical method can be used.

The solution of the equation $\left|[A] - \lambda[I]\right| = 0$ yields the three values of λ: $\lambda_1 = 0$, $\lambda_2 = -1$ and $\lambda_3 = -2$.

The solution of the matrix equation:

$$[A - \lambda_i\ I]\ \{P_i\} = 0 \qquad (2.16)$$

where the vector $\left\{P_i\right\} = \begin{Bmatrix} P_{i1} \\ P_{i2} \\ P_{i3} \end{Bmatrix}$

is an eigenvector associated with the eigenvalue λ_i for i = 1, 2 and 3 will give the three columns of the matrix $\left[P\right]$.

With i = 1, $\lambda_1 = 0$ gives

$$\begin{aligned} P_{12} &= 0 \\ P_{13} &= 0 \\ -2P_{12} - 3P_{13} &= 0 \end{aligned} \tag{2.17}$$

This yields

$$\left\{P_1\right\} = \begin{Bmatrix} a \\ 0 \\ 0 \end{Bmatrix}$$

where a is an arbitrary real number.

Then i = 2, $\lambda_2 = -1$ gives

$$\begin{aligned} P_{21} + P_{22} \qquad &= 0 \\ P_{22} + P_{23} &= 0 \\ -2\,P_{22} - 2\,P_{23} &= 0 \end{aligned} \tag{2.18}$$

This yields

$$\left\{P_2\right\} = \begin{Bmatrix} b \\ -b \\ b \end{Bmatrix}$$

where b is an arbitrary real number.

Then i = 3, $\lambda_3 = -2$ gives

$$\begin{aligned} 2P_{31} + P_{32} \qquad &= 0 \\ 2P_{32} + P_{33} &= 0 \\ -2P_{32} - 1P_{33} &= 0 \end{aligned} \tag{2.19}$$

This yields

$$\{P_3\} = \begin{Bmatrix} c/4 \\ -c/2 \\ c \end{Bmatrix}$$

where c is an arbitrary real number.

The matrix is therefore given below together with the inverse matrix:

$$[P] = \begin{bmatrix} a & b & c/4 \\ 0 & -b & -c/2 \\ 0 & b & c \end{bmatrix} \qquad [P]^{-1} = \begin{bmatrix} 1/a & \frac{3}{2a} & \frac{1}{2a} \\ 0 & -\frac{2}{b} & -1/b \\ 0 & 2/c & -2/c \end{bmatrix}$$

Substituting the transformation $\{e\} = [P]\{y\}$ into equation (2.12)

$$[P]\{\dot{y}\} = [A][P]\{y\} + [B]\{u\} + \{f\}$$

$$\therefore \quad \{\dot{y}\} = [P]^{-1}[A][P]\{y\} + [P]^{-1}[B]\{u\} + [P]^{-1}\{f\} \qquad (2.20)$$

$$\therefore \quad \{\dot{y}\} = [\Lambda]\, y + [W]\,\{u\} + [P]^{-1}\,\{f\}$$

For the example $[\Lambda]$ can now be calculated easily as:

$$[\Lambda] = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$ which is independent of the constants a, b, and c.

and choosing for convenience $a = -\frac{1}{2}$, $b = 1$ and $c = 2$ then:

$$[W] = \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} \quad \text{and} \quad [P]^{-1}\{f\} = [Q]\{f\} = \begin{Bmatrix} -f \\ -f \\ -f \end{Bmatrix}$$

In terms of the new variables, the canonic state variables $\{y\}$, equation (2.11) can be written:

$$\begin{Bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{Bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{Bmatrix} y_1 \\ y_2 \\ y_3 \end{Bmatrix} + \begin{Bmatrix} 1 \\ 1 \\ 1 \end{Bmatrix} u + \begin{Bmatrix} -f \\ -f \\ -f \end{Bmatrix} \tag{2.21}$$

This is a canonical form for the original equation which is convenient mathematically as the variables are separated and the constants have been reduced to unity.

Consider a component equation of the last form of equation (2.20):

$$\dot{y}_i = \lambda_i y_i + \sum_k W_{ik} u_k + \sum_j Q_{ij} f_j \tag{2.22}$$

This equation is integrable if the functions u_i and f_i are real and measureable and if the initial condition vector $\{y(0)\}$ is known.
Multiply (2.21) by $e^{-\lambda_i t}$ then:

$$\frac{d}{dt}\left(y_i e^{-\lambda_i t}\right) = e^{-\lambda_i t} \sum_k W_{ik} u_k + e^{-\lambda_i t} \sum_j Q_{ij} f_j \tag{2.23}$$

and

$$y_i = e^{\lambda_i t}\{y(0)\} + e^{\lambda_i t}\int_0^t e^{-\lambda_i \tau} \sum_k W_{ik} u_k d\tau +$$

$$+ e^{\lambda_i t}\int_0^t e^{-\lambda_i \tau} \sum_j Q_{ij} f_j d\tau \tag{2.24}$$

or, returning to matrix notation, the general solution is:

$$\{y\} = [G]\{y(0)\} + [G]\int_0^t [G]^{-1}[P]^{-1}[B]\{u\}\, d\tau + [G]\int_0^t [G]^{-1}[P]^{-1}\{f\}\, d\tau \tag{2.25}$$

The matrix $[G]$ is called the "system impulse response matrix" and is defined:

$$[G] = e^{-[\Lambda]t} = [I] + [\Lambda]t + \frac{[\Lambda]^2 t^2}{2!} \ldots\ldots\ldots \qquad (2.26)$$

for linear systems in canonical form.

Transform equation (2.24) back to the original variables using $\{y\} = [P]^{-1}\{w\}$:

$$[P]^{-1}\{e\} = [G][P]^{-1}\{e(0)\} + [G]\int_0^t [G]^{-1}[P]^{-1}[B]\{u\}d\tau + [G]\int_0^t [G]^{-1}[P]^{-1}\{f\}d\tau \,. \qquad (2.27)$$

and multiply by $[P]$ and introduce $[P]^{-1}[P]$ into the integral terms:

$$\{e\} = [P][G][P]^{-1}\{e(0)\} + [P][G][P]^{-1}\int_0^t [P][G]^{-1}[P]^{-1}[B]\{u\}\,d\tau + [P][G][P]^{-1}\int_0^t [P][G]^{-1}[P]^{-1}\{f\}\,d\tau \qquad (2.28)$$

Now writing $[H] = [P][G][P]^{-1}$ and therefore

$$[H]^{-1} = [P][G]^{-1}[P]^{-1} \qquad (2.29)$$

equation (2.29) becomes:

$$\{e\} = [H]\{e(0)\} + [H]\int_0^t [H]^{-1}[B]\{u\}d\tau + [H]\int_0^t [H]^{-1}\{f\}\,d\tau \qquad (2.30)$$

The matrix $[H]$ is now the system impulse response matrix in terms of the original variables $e_1 \ldots\ldots e_n$ and is defined in terms of the system matrix $[A]$:

$$[H] = e^{[A]t} = [I] + [A]t + \frac{[A]^2 t^2}{2!} \ldots\ldots \qquad (2.31)$$

A system impulse response matrix $[H]$ can always be obtained from knowledge of the system matrix $[A]$. The system matrix $[A]$ cannot always be diagonalized, to yield the matrix $[\Lambda]$, however, unless the eigenvalues are real, and distinct. The matrix $[A]$ can always be put into the Jordan

canonical form, using a suitable transformation ([24], p. 287) ([21], p. 191) and the resulting equations in terms of canonic state variables can be integrated.

Returning to the example of this section for a moment and writing equation (2.21) in component form gives the three equations:

$$\begin{aligned} \dot{y}_1 &= (0)y_1 + u - f \\ \dot{y}_2 &= (-1)y_2 + u - f \\ \dot{y}_3 &= (-2)y_3 + u - f \end{aligned} \tag{2.32}$$

and Laplace transforming:

$$\begin{aligned} sY_1 - (0)Y_1 &= U - F \\ sY_2 - (-1)Y_2 &= U - F \\ sY_3 - (-2)Y_3 &= U - F \end{aligned} \tag{2.33}$$

The block diagram of the equations (2.33) is drawn in Figure 2.3. The fact that the variables have been "separated" can be seen clearly by comparing Fig. 2.3 with the original block diagram (Fig. 2.1).

The systems to be discussed in the remainder of this volume will frequently be expressed in terms of canonic state variables.

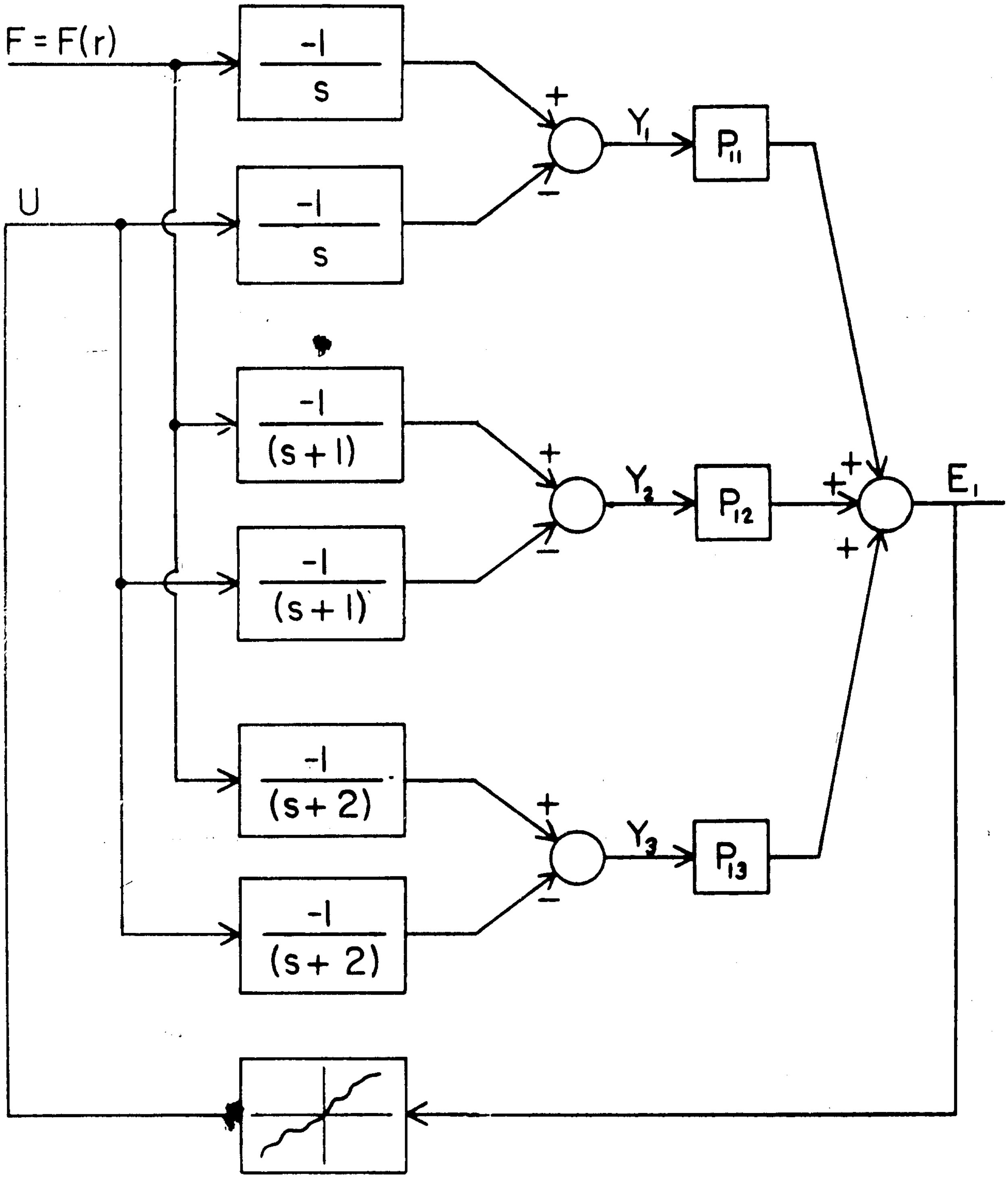

Figure 2.3

The System of Figure 2.1 Redrawn in

Terms of the Canonic State Variables

2.5 The Second Method of Liapunov

The Second (Direct) Method Of Liapunov (SML) is theoretically the most general available method for stability analysis of nonlinear systems. A detailed mathematical discussion of the SML is contained in the books by Malkin [25], Zubov [26], Hahn [27] and in various papers, notably those by Kalman and Bertram [28] and La Salle [7]. An introductory treatment of the SML and some of its engineering applications are contained in [29] (Boston Workshop on the SML). Technical Report TR-61-6 of this contract [30] deals with the engineering applications of Liapunov's second method.

Three major limitations of the SML in the analysis of autonomous nonlinear physical systems are presently:

1. There are no known straight-forward procedures of constructing Liapunov functions for the general class of nonlinear autonomous systems. One's success depends largely upon intuition and experience.
2. The known Liapunov functions for special types of nonlinear systems yield sufficient but not necessary conditions for stability.
3. The SML is, at the present state of the art, not directly applicable to systems with limit cycles, no matter how small and insignificant the limit cycle oscillations may be.

A survey of the most widely applicable methods of constructing Liapunov functions, including some results of research at Purdue, is contained in Technical Report TR-61-6 of this project [30]. Many autonomous systems containing nonlinear gain elements can be analyzed successfully by the SML by

means of the canonic transformations of Lur'e [31], Letov [32] and the pseudo-canonic transformations developed at Purdue [35]. Attempts have recently been reported to analyze, by the SML, the stability of relay (switched) systems (Alimov, [33]) and systems with time delay (transportation lag), (Razumikin, [34]).

The failure of the SML to yield necessary conditions for stability is frequently the result of its inability to predict limit cycle oscillations. Some progress in extending the applicability of the SML to systems containing limit cycles has been reported by Zubov [26] and La Salle [7]. Rekasius and Szego developed a procedure whereby one is able to find a closed, bounded region in the state space in which the limit cycle is confined, without the need for exact solution of the limit cycle [35].

Hence the present day practical limitations in the applicability of the SML in stability analysis of autonomous nonlinear systems are gradually diminishing. It appears that continued research efforts will make the SML a very practical and powerful tool for the stability analysis of autonomous nonlinear systems.

CHAPTER 3

STABILITY OF NON-AUTONOMOUS SYSTEMS

3.1 Introduction

Despite the fact that autonomous and nonautonomous systems have been defined precisely in Chapter 1 of this volume, it will be worthwhile to review quickly those definitions and discuss their applicability in this chapter. The term "autonomous" refers to a free (unforced) time invariant system whereas the term "nonautonomous" refers to a time invariant (stationary) system subjected to inputs (forced system) or to time variable parameter (nonstationary) systems irrespective of whether they are forced or not. In this chapter a distinction between unforced and forced will be made instead of a distinction between autonomous and nonautonomous systems.

It will suffice to mention at this point that the problem of determining the stability of a nonstationary nonlinear forced system should be relegated into the background, until the problem of obtaining the stability information of a stationary, nonlinear forced system is solved.

Considerable effort has been expended by various researchers, particularly by mathematicians investigating the stability theory of differential equations, to obtain methods of determining the stability of unforced systems. In general, an unforced system is a fiction which does not exist in practice. Every control system is forced, either due to inputs or disturbances or both.

One possible reason for the existence and continuing increase of the vast amount of literature dealing with the stability of nonlinear unforced

systems by technical journals may be due to the fact that most engineers still think of nonlinear systems in terms of analogous linear time-invariant systems. It is a fairly common practice to try to extend familiar concepts applicable to special cases to more general cases. Unfortunately, this method often leads nowhere. This is evidenced, for example, by the tremendous though essentially unsuccessful efforts that have been made to extend the use of the familiar Laplace and Fourier transforms to analyze linear time variable parameter systems [23].

The stability characteristics of a linear system are the same irrespective of whether there are any inputs to the system or not. Hence it is common practice while studying the stability of linear systems to consider only the unforced case. There is considerable justification in adopting this procedure since the stability of both the forced and unforced systems are simultaneously determined.

A practicing control engineer has very little use for methods which yield stability information for unforced systems only, since every actual control system is governed by a differential equation with a forcing function. Unfortunately, most methods that are available at the present to investigate the stability of nonlinear systems seem to be applicable only to the unforced case. Even a regulator is not an unforced system since, despite the fact that the input is a constant and hence the deviations of the input from a steady state value are zero, the output and load disturbances make the system forced.

The last paragraph should not be interpreted to mean that the stability of the unforced system is unimportant. It is quite possible, however, that an unstable (in the sense that limit cycles of undesirable amplitudes might exist in the system) unforced system may become stable (in the sense that

the limit cycle may be reduced in amplitude or quenched altogether) when subjected to inputs. A special case of this occurrence is the phenomenon of signal stabilization, discussed later. However, it is also quite likely that for some period of time the system may be exposed to constant inputs or load disturbances. In this case the system is mathematically equivalent to an unforced system. Hence it is necessary to impose restrictions on the stability characteristics of the unforced system. The comments in the last paragraph apply to methods which are useful for investigating unforced systems only and not to the unforced systems.

The nonexistence of suitable methods for investigating the stability of forced nonlinear systems is further complicated by the very concept of stability for these systems. The familiar concept of stability which is straightforward and intuitively easy to understand in the case of linear time invariant systems takes on a more subtle and difficult aspect in the case of nonlinear autonomous systems in general and nonlinear nonautonomous systems in particular. Antosiewicz [36] defines several distinctly different types of stability for nonlinear systems.

Considerable research is warranted before any conclusions may be drawn regarding methods of investigating stability of forced nonlinear systems. One general method which is capable of further extension and two special inter-related methods useful for investigating the stability of certain specific stationary nonlinear forced systems are considered in this section. Needless to say, the philosophy of presentation of this section may seem to have overtones of pessimism because of the present state of the art of nonlinear systems in general and nonlinear forced systems in particular.

3.2 The Second Method of Liapunov

While the SML still is, theoretically, the most general available method for stability analysis of unforced nonlinear systems, its practical application is still in its infancy despite the fact that several special techniques are available for specific nonlinear unforced systems. To emphasize the enormous difficulties encountered in stability analysis of unforced nonlinear systems, it is sufficient to note that even the problem of linear time-varying systems still awaits its solution.

As pointed out earlier, additional difficulties encountered in the application of the SML to forced systems are due to a number of distinctly different types of stability which manifest themselves only in nonautonomous systems. Consequently the theorems of the SML of stability and instability take on different forms, depending upon the type of stability which is to be proved. Many stability and instability theorems for unforced systems, stationary and nonstationary, based upon the SMl are contained in the books by Hahn [21], Zubov [26], Malkin [25] and in the papers of Antosiewicz [36] and Kalman and Bertram [28]. Very little is known, however, at the present time of how to construct Liapunov functions for nonautonomous systems. A few studies of stability of special cases of time-varying parameter systems are scattered in the periodical literature, primarily in various issues of Automatika i Telemekanika (Automation and Remote Control) and Prikladnaja Matematika e Mekanika, (P.M.M.).

While there is little hope yet for a major breakthrough in the practical application of the SML and the methods of construction of Liapunov functions for the general case of forced nonlinear system, some special cases may in the near future become practically managable. These are,

for example, the stability of linear time varying systems (Szego, [37]), and the analysis of systems with periodically varying coefficients, etc. Despite the fact that the solution of this problem does not solve the problem of determining the stability of nonlinear forced systems, it is hoped that it will provide some insight into the latter problem.

Very little is known about the problem of the stability of a general nonlinear system subjected to inputs from the point of view of the SML. However it is sometimes possible to invoke Massera's theorem [38] which, in essence, states that a sufficient condition for the total stability of a forced nonlinear system (stationary or nonstationary) is that the unforced system be uniformly asymptotically stable. Massera's theorem is still not very useful for nonstationary nonlinear systems since, as pointed out earlier, the application of the SML even to nonstationary linear systems is not easy. However, Massera's theorem may have some use in the case of a stationary nonlinear system since certain methods for applying the SML to certain special classes of nonlinear systems are available in the literature. Notice, however, that the use of Massera's theorem imposes severe restrictions on the stability characteristics of the unforced system. Uniform asymptotic stability may be a sufficient but not necessary condition for acceptance of an engineering system. This condition excludes, for example, all systems which may possess small limit cycles for some specific values of the system parameters.

At the present state of the art the SML for forced nonlinear systems is a fruitful area of research but has so far yielded very little of practical importance.

3.3 Signal Stabilization

The stability characteristics of a linear system are unaffected by the inputs to the system. This however is not true, in general, for nonlinear systems. The possibility of changing the stability characteristics with different inputs is the property which allows "signal stabilization".

Feedback Control Systems in a state of self sustained oscillations (limit cycle operation) resulting in output hunt may often be stabilized by the introduction of an external signal of a sufficiently high frequency at a convenient point in the loop. This phenomenon is termed "signal stabilization" by Oldenburger [39]. Here a system is said to be stabilized if the amplitude of the output hunt is reduced below a certain prescribed value. A first attempt to explain this phenomenon when the waveform of the "stabilizing signal" is sinusoidal is due to Oldenburger and Liu [40]. The theory developed by Oldenburger and Liu is quite different from the one advanced by Minorsky [41], who treated the use of a signal to excite or quench the hunt (self oscillation) of a physical system described by a particular type of second order differential equation. Oldenburger and Nakada [42] extend the theory of signal stabilization to a rather general class of nonlinear systems with a triangular waveform stabilizing signal. Sridhar and Oldenburger [43], [44] generalize the theory of signal stabilization and extend it to consider random stabilizing signals. They also establish various criteria to obtain stability information for a particular class of nonlinear systems. Oldenburger and Boyer [45] generalize the theory developed in reference [40] for sinusoidal stabilizing signals.

Signal stabilization theory as developed in references [40] and [41] to [46] appears to hinge on the fact that the frequency of every component

in the stabilizing signal is large compared to the significant frequencies in the system. This assumption is consistent with the practical use of signal stabilization for decreasing the output hunt in a self-oscillating system, since it is desired that neither the system hunt nor the stabilizing signal be present to any appreciable degree at the output. However, the theory developed in reference [44] may easily be extended to cover the case when the input spectrum has low frequency components.

Recently Gibson and Sridhar [46] have proposed a new method for considering certain specific nonlinear systems with sinusoidal inputs without putting any restrictions on the frequency of the input. This method will be discussed further in the next section.

It is felt that the theory of signal stabilization provides a better insight into the problem of understanding the stability characteristics of a particular class of forced nonlinear systems. It should be pointed out that it may be possible to interpret a signal stabilized nonlinear system as either a forced or unforced system, depending on whether the stabilizing signal generator is included within the "black box" representing the nonlinear system or not.

3.4 The Dual Input Describing Function

The describing function (D.F.) is a very useful approximation in the analysis of a certain class of nonlinear systems. It applies directly to systems such as those shown in Fig. 3.1. It is based on the method of harmonic balance of Kryloff and Bogoliuboff [8] and, as discussed above, was applied to control systems by Goldfarb [9]. Popov [47] has an interesting discussion of the method of harmonic balance itself as it applies to control systems. In all of this work the system under analysis is un-

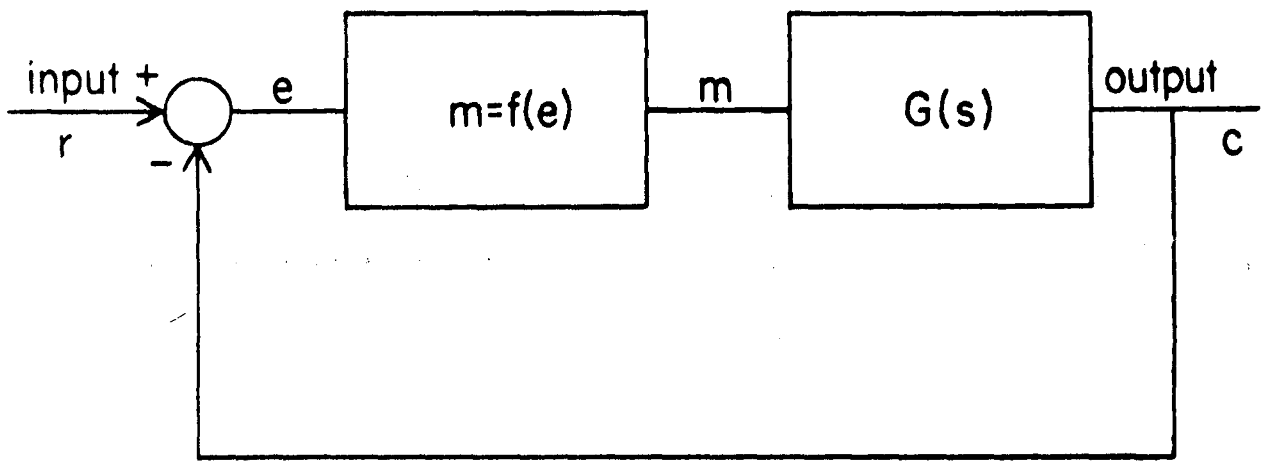

Fig. 3.1

The Type of System for Which the Describing Function Method of Analysis is Applicable.

forced. It seems a direct step, however, to apply the conventional D.F. to forced systems.

A number of schemes have been proposed for obtaining the closed loop frequency response of a nonlinear system, for example, by direct extension of the conventional D.F. . Among these methods, those of Levinson [15], Thaler [48] and Ogata [49] are well known. Hill [16] has proposed an ingenious use of the Nichols chart which is probably the most convenient of all of these techniques. Kochenburger [10] in his original paper discussed the extension of M peak to the D.F. plot and presumes that one can read off the amplitude of the resonant peak of a sinusoidally driven nonlinear system from the D.F. plot just as one does from the Nyquist plot for a linear system. Prince [50] has proposed a modification of the conventional D.F. to obtain the closed loop response of a perfect relay system. However the Prince D.F. does not appear to be of wide applicability.

The error in all of the work cited above lies in the fact that the conventional D.F. analysis postulates a single sinusoidal input to the nonlinear elements. Naturally the frequency chosen will be that of the input r. Now if the closed loop system is (uniformly) asymptotically stable, then with an input signal, this analysis is as valid as the conventional D.F. analysis of unforced systems. This is so because in fact there will be the single sinusoidal signal at e, for which the conventional D.F. analysis is designed.

Suppose, however, that the control system is not asymptotically stable in the presence of the input r. Then the conventional D.F. is in error because there is no longer a single sinusoidal signal at e. It is a well

known fact that in a nonlinear system asymptotic stability or instability of the unforced case doe not imply either stability or instability of the forced system. Therefore it can be concluded that it is improper to employ conventional D.F. for closed loop response calculations unless the stability of the driven system has been established by other means. This fact is apparently not appreciated by a significant segment of control engineers.

A number of dual input describing functions (DIDF) have been proposed. However, they may be applied to closed loop frequency response calculations only under certain conditions that do not usually hold. West, Douce and Livesly [51] have proposed a DIDF that is valid only if the two sinusoidal components at the input to the nonlinearity are related by an integer. This DIDF is rather clumsy to manipulate, but it can be used to detect subharmonic response. It cannot be used to examine the general possibility of asynchronous oscillations induced by the input in general, however. Oldenburger and Boyer [45] have proposed a DIDF that is more convenient to manipulate, but that is valid only if the two sine waves at the input to the nonlinearity are widely separated in frequency. Thus this approach is useless within the bandpass of the system. Sridhar and Oldenburger Ref [43],[44] have developed a DIDF in which one of the signals at e is a stationary, Gaussian, random function. It appears that this function may be employed to develop the response of a nonlinear system to a random input. This problem has been considered by Booton 52 , but of course the same objection a to previous work with sine waves applies to this: it completely ignores the stability problem. Gibson and Sridhar [46] have applied a general DIDF developed by Sridhar [53] to the problem of closed loop frequency response and interesting results have been obtained. It is shown in reference [46] that stable

unforced systems may become unstable under certain driving functions and that also the converse is true.

It is apparent that the DIDF must be developed until it is as simple and reliable for forced systems as the conventional DF is for unforced systems if it is to be useful for the specification of automatic control systems for aero space vehicles. With the present rapid rate of research progress in this area, it is possible that this will occur within the next few years.

3.5 Conclusions

It is hoped that this chapter will throw some light onto the magnitude of the problem involved in considering the stability of forced systems. Considerable research on the problem of determining practical methods for obtaining the stability of forced nonlinear systems must be conducted before any significent progress can be reported in this area. Even the discovery of some approximate methods for determining the stability of certain classes of forced nonlinear systems, such as the describing function method for a special class of unforced nonlinear systems, would be a definite contribution. It does not appear at the moment that a unified method of stability analysis applicable to all forced nonlinear systems will be discovered in the foreseeable future, if at all. This last statement appears to be reasonable in the light of the trend experienced in the field of nonlinear mechanics, where a number of special methods for obtaining stability and other properties of a small number of special classes of systems is available. This same approach of trying to obtain special methods for different types of forced nonlinear control systems is being adopted at the present.

Despite the fact that most specifications that might eventually be

recommended for nonlinear systems may involve the response of the system to specific inputs, it is felt that the problem of stability of the forced system is intimately related to its response to inputs. Thus, for example, it is possible to have a "stability specification" which states that a limit cycle amplitude larger then a certain value cannot be tolerated. The specified amplitude, of course, will depend on the applications.

It is felt that with the present state of the art, most of the research effort for determining the stability of forced nonlinear systems should be concentrated on obtaining methods for determining this information for stationary systems. It is hoped that solutions to this problem will pave the way for better understanding of the problem and eventual solution of the stability of forced nonstationary systems.

CHAPTER IV

THE RESPONSE OF AUTONOMOUS SYSTEMS

4.1 Introduction

A possible approach to the problem of specifications for nonlinear systems is the construction of a mathematical model which is representative of the best system that can be devised for a given task. This system, which is optimum with respect to certain specific requirements, and its performance can be used as the upper bound on physical, but not necessarily optimum, systems.

The question of which model is optimum for a given task must include, in general, consideration of such qualities as reliability, economy and performance, to quote three examples. In addition, one engineers' optimum may well differ from another engineers' optimum within a given task.

The problem has been formulated in the literature (Bellman [54], p. 22), (Merriam [55], p. 267), (Lee [56]) in terms of a classical problem in the calculus of variations (Forsyth [57], Chapter 1). Here an index of performance, $J(x,y)$, is to be minimized (or maximized) by choice of the function y:

$$J(x,y) = \int_0^T k(x, y)dt \tag{4-1}$$

where the vector $\{x\}$ represents the system state variables

the vector $\{y\}$ is the system steering function

and the time T is related to the termination of the control problem. The function k is chosen to include the considerations mentioned above and the constraints of a given problem. In practice the choice of this

function usually involves a compromise between an accurate evaluation of the physical process and a more tractable mathematical problem.

Solution for the function y as a function of time then defines an optimum policy for the system, should such a policy exist, by means of which optimum performance is achieved. General problems of this nature are frequently insolvable.

The purpose of this chapter will be to examine, therefore, a specialization within the general problem which has received attention in the literature. The class of system to be considered are those that are autonomous ([1], p. 32) and where y is to be determined so that the disturbed response is time optimum.

The study of this restricted class of systems together with the restricted nature of the performance index is warranted as it permits exploration of the techniques useful with these rather difficult problems. The chapter reflects the state of the art and indicates that the approach has much promise, but that there is the need for further work in this area.

4.2 Time Optimum Switched Systems

Engineers always try to build the best system possible from every point of view, e.g., reliability, economy and performance. Often one system quality must be sacrificed for another, and the resulting system is then the best that can be built, i.e., optimum, after having taken all factors into consideration. A specialization within this optimum concept is the performance specification of being "time optimum". The question to be answered here is how should a system be built so that it will achieve its objectives in minimum time.

Some time ago engineers began to reason that perhaps the system that

could use the maximum power available, all of the time, would be time optimum. This idea is contrary to the concept of a linear system where the maximum power available is used only for one instant of time and a lesser amount used at all other times. The intuitive conclusion at this stage was that a relay system and a time optimum system were one and the same thing.

A relay system is a nonlinear system with a fundamental property that the nonlinearity, the relay, is separable from the linear portion of the system. The configuration is like that of Figure 4.1 , rather than the linear system shown in Figure 4.2.

Early attempts to analyse such systems were restricted to cases where the linear portion of the system had a relatively simple form, frequently

$$G(s) = \frac{1}{s^2} \text{ or } G(s) = \frac{1}{s(1+Ts)}$$

(Bogner [58]), (Oldenburger [59]), (Weiswander [60]),(Kahn [61]).

Once the relay has been included in the circuit the question arises, when must it switch? If the phase plane is used (this representation is applicable to the second order examples of the paragraph above), where do the switching boundaries lie?

The system shown in Figure 4.1 will switch along a line that is the ordinate axis, Figure 4.3. Here the objective would be to reduce the error and the derivative of the error to zero. Variations of the switching boundaries include linear switching, Figure 4.4, and parabolic switching, Figure 4.5. In each of these three diagrams there are several possibilities; for example, changing the switching to a different quadrant of the phase plane, or interchanging the relay polarity.

With the introduction of the more complicated switching boundary, an

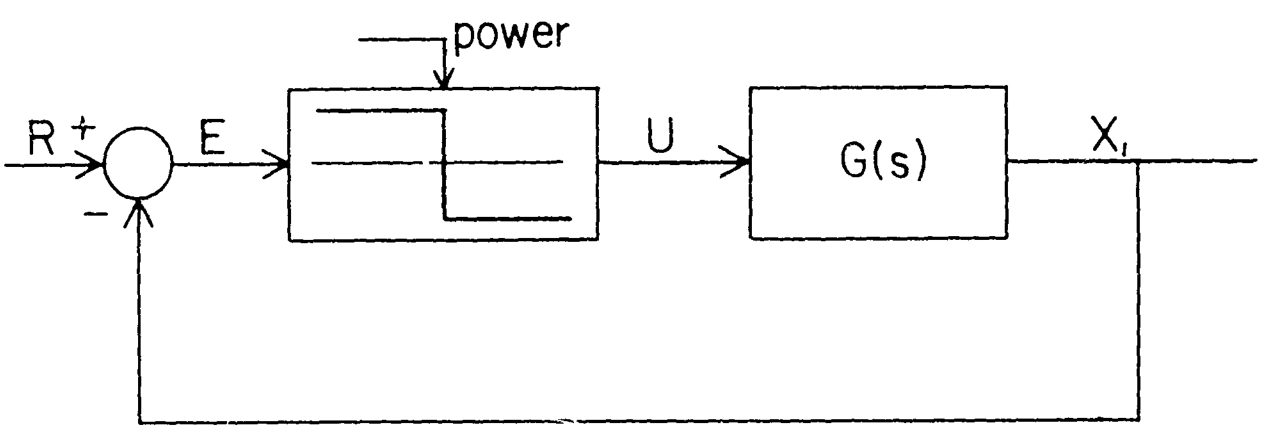

Figure 4.1 System with a Separable Nonlinearity

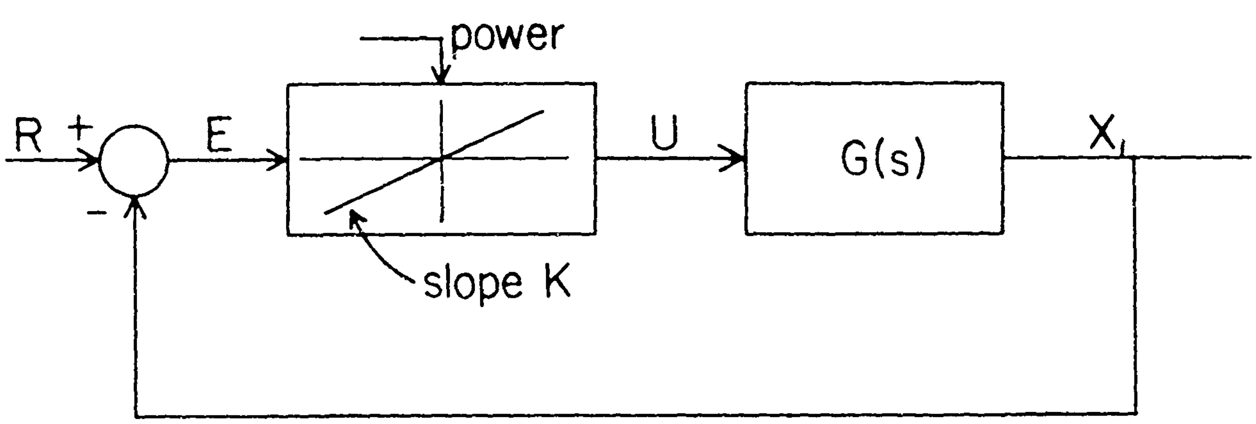

Figure 4.2 Linear System with Gain K

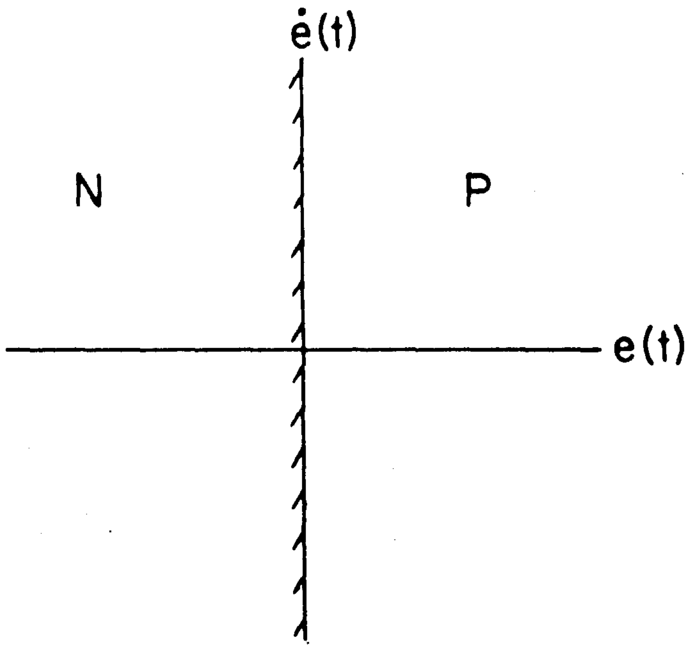

Figure 4.3

Simple Switching Boundary for the System of Figure 4.1

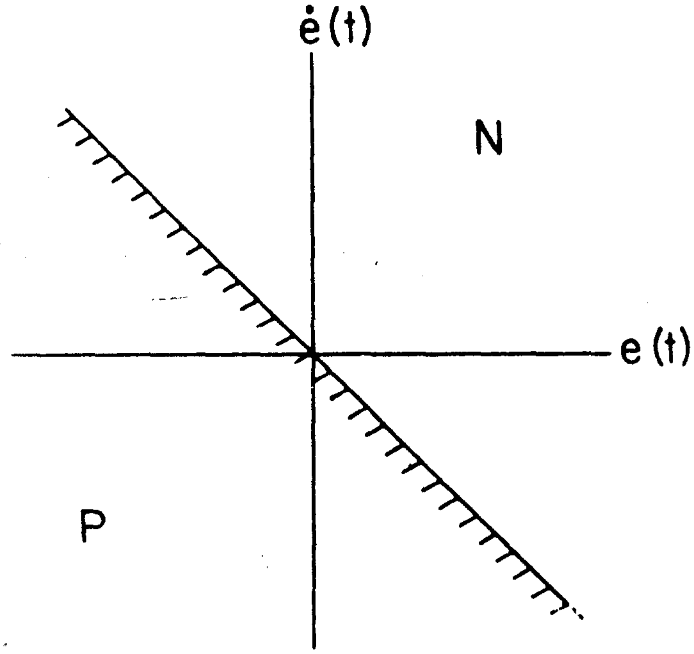

Figure 4.4

Linear Switching Boundary

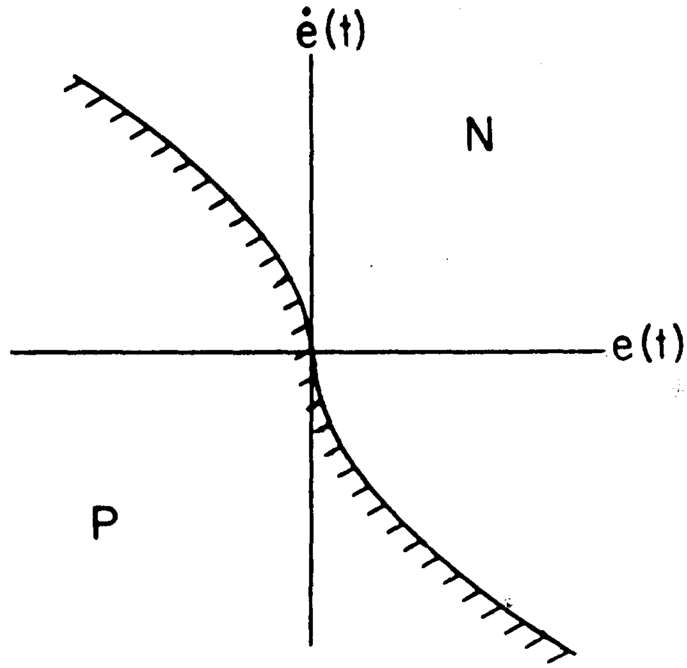

Figure 4.5

Parabolic Switching Boundary

additional element must be added to the system. This element will have the task of determining where the system variables are with respect to the boundary and which polarity to feed to the relay. The element will be a form of computer and the configuration becomes that of Figure 4.6 for the system corresponding to Figure 4.4 or 4.5.

None of the systems mentioned yet could be called successful, however, except in restrictive cases. For example the configuration of Figure 4.1 with the boundary of Figure 4.3 will switch many times before a region near the origin is reached, and then it will oscillate about the origin (limit cycle). With the switching boundary of Figure 4.4, the system will only reach the vicinity of the origin from a discrete number of points on each side of the boundary. From all other points the system will drive toward one of two points on the abcissa, on either side of the origin. The points will correspond to the magnitude of the relay output. Neither of these cases are time optimum nor are they optimum in any sense.

The parabolic boundary of Figure 4.5 is time optimum for the specialized system with the linear portion described by $G(s) = 1/S^2$. The technique used to deduce this boundary ([58], p. 117) is not suitable for use with other systems as it depends on the phase plane technique and the restricted nature of the system considered.

The relay used has so far been considered ideal. No relay is ideal and a number of authors have attempted to extend the techniques used with these rather special systems and boundaries to allow for physical relay characteristics such as deadband and hysteresis [60], (Izawa [61]).

It is to be noted that any physical system must have deadband in the relay mechanism in order to deactivate the system when it reaches the origin.

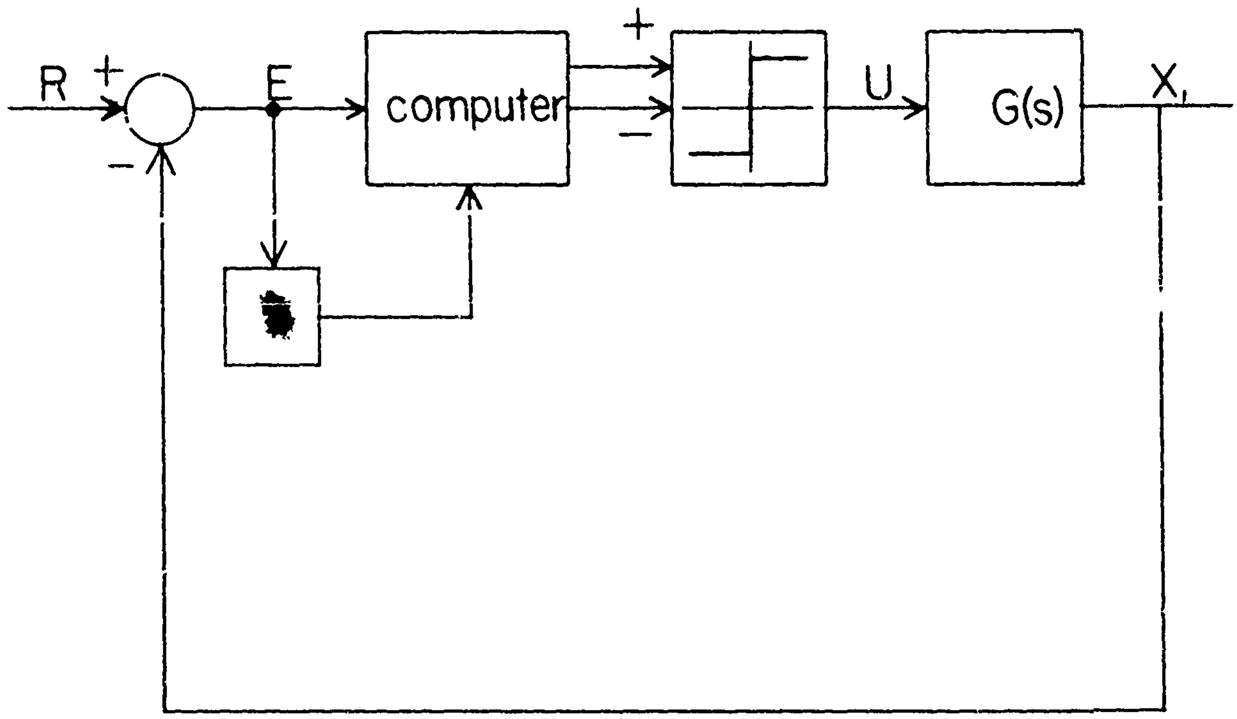

Figure 4.6

System Configuration for Nonsimple

Switching Boundaries

Alternatively the computing element can be designed to allow the system to operate linearly in a region near the origin. This is the dual-mode system (McDonald [63]), (Bulund [64]), a simple example of which can be derived from Figure 4.4 as shown in Figure 4.7. The boundary of the linear region of operation can take a number of forms and is left undefined in Figure 4.7 for this reason. Some such procedure of deactivation or restricted linear operation is necessary with all practical systems.

It was suggested at this stage of the development of optimum switched systems that the number of switchings needed for systems with real, distinct roots, associated with the linear portion, is (n - 1), where n is the system order (discussion to [60]), [58].

The next development in the state of the art was the complete analysis of a second order system, again with an ideal relay. The systems investigated were those that could be described by equations of the form:

$$\ddot{y} + 2\zeta\dot{y} + y = \pm 1 \qquad (4\text{-}2)$$

The investigators sought out every possible mode of operation and by systematic elimination converged on the optimum (Flugge - Lotz [65]), (Bushaw [66]), (Tsein [67], p. 136). Sufficient theorems and lemmas were proven to substantiate the elimination process and the optimum was proven optimum. The results reported by Bushaw in his Ph.D. thesis [66] and reproduced by Tsein [67] show, for example, that for the case where $\zeta = 0$ in equation (4.1) the switching boundaries for optimum time response are portions of the circles associated with the system singular points in the phase plane, centers in this case, as in Figure 4.8.

It is fairly obvious that the tool whereby the results obtained so far had been obtained is the phase plane method and the geometrical interpreta-

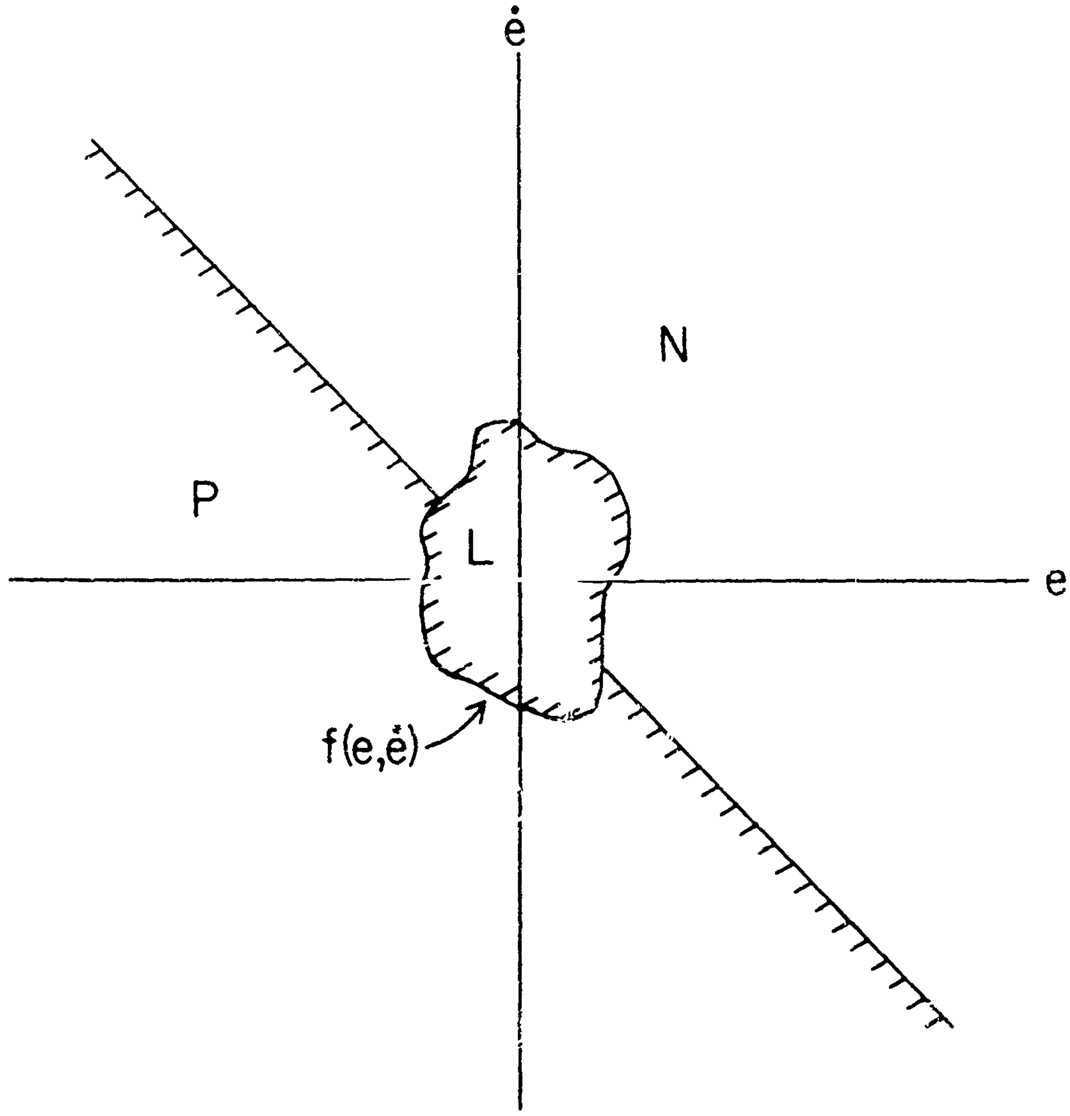

Figure 4.7

Switching Boundaries for a

Dual Mode System

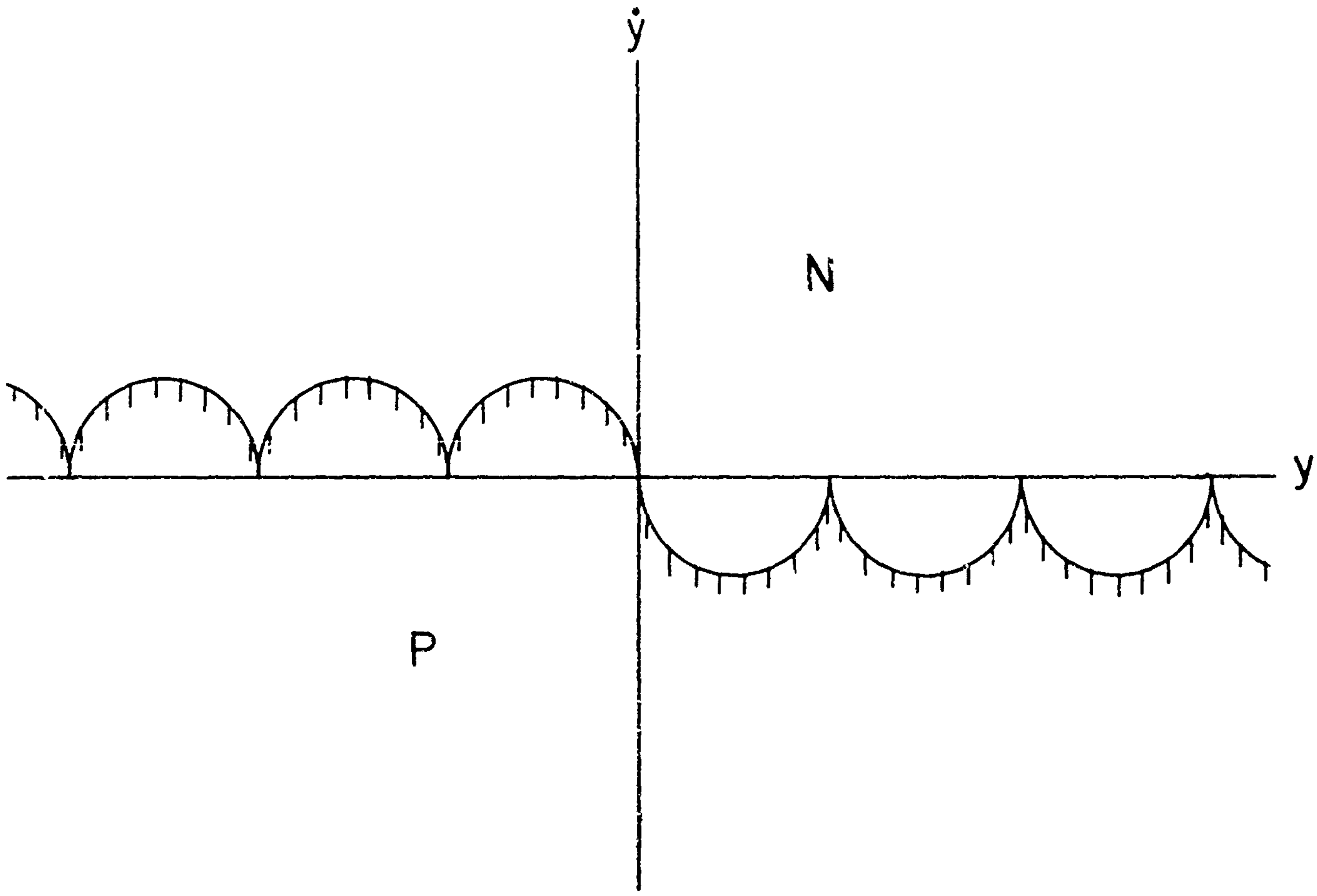

Figure 4.8

Optimum Switching Boundaries for a

Second Order System with Zero Damping

tion possible there. Third order system analysis has been attempted on the phase plane, or rather on two phase planes [54], (Chang [68]). The method is, however, rather cumbersome and for practical purposes the phase plane is restricted to second order systems. This limitation has led to the use of a state space and state variables (described elsewhere in this volume) and the use of more elegant mathematics.

Consider the configuration shown in Figure 4.9 which is the forward transfer function of the system to be examined. The nonlinearity is not defined, as it is the variable to be used in the time optimization process. It is, however, constrained to be a real, measurable function of the input variable m and bounded above and below such that $-1 \leq |u_1(t)| \leq 1$. The form of the linear portion of the system is not necessarily constrained to the form shown in Figure 4.9. The form can vary considerably with the only restriction that the system equations can be written in matrix notation and in canonical form as described elsewhere in this volume (Chapter 2).

The system equations of this example are:

$$\begin{Bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{Bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -3 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ 1 \end{Bmatrix} u + \begin{Bmatrix} 0 \\ 0 \\ f \end{Bmatrix} \tag{4-3}$$

or ([20] , equation 1)

$$\{\dot{x}\} = [A]\{x\} + [B]\{u\} + \{f\} \tag{4-4}$$

Transforming from the physical variables $\{x\}$ to the state variables $\{y\}$ with the transformation, $\{X\}$ is:

$$\{x\} = [H]\{y\} \tag{4-5}$$

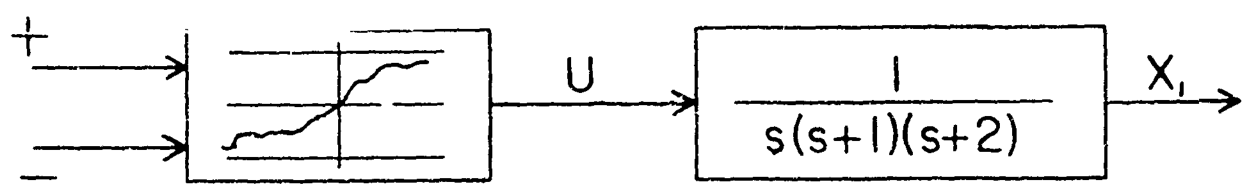

Figure 4.9

Forward Transfer Function of a System

with a Separable but Undefined Nonlinearity

where the matrix $[H]$ is given by a solution of the matrix equation:

$$[\dot{H}] = [A][H] \text{ with initial conditions } [H(0)] = [I] \; . \qquad (4\text{-}6)$$

The matrix $[H]$ is defined as the system impulse response matrix in terms of the variables $(x_1 \ldots x_n)$ and may also be determined as:

$$[H] = e^{[A]t} \qquad (4\text{-}7)$$

where $[A]$ is the system matrix.

The solution to equation (4-4) may now be determined by Lagranges method of variation of parameters ([21], chapter 10, section 12). Differentiating (4-5) and substituting in (4-4):

$$\begin{aligned} \{\dot{x}\} &= [\dot{H}]\{y\} + [H]\{\dot{y}\} \\ \{\dot{x}\} &= [A][H]\{y\} + [B]\{u\} + \{f\} \end{aligned} \qquad (4\text{-}8)$$

Comparing these two equations with the help of (4-6) one gets:

$$[H]\{\dot{y}\} = [B]\{u\} + \{f\} \qquad (4\text{-}9)$$

Integrating, returning to the original variables and using the initial condition vector $\{x(0)\}$, $\{x\}$ is found to be:

$$\{x(t)\} = [H]\{x(0)\} + [H]\int_0^t [H]^{-1}[B]\{u\}\, d\tau + [H]\int_0^t [H]^{-1}\{f\}\, d\tau \qquad (4\text{-}10)$$

This is the solution of the system equations and can be found provided equation (4-6) can be solved and the integrations can be carried out. The solution of equation (4-6) is unfortunately a difficult if not impossible task in the general case. Whether or not these integrations can be performed depends largely on the form of $\{u(t)\}$. The restrictions already placed upon the nonlinearity will, however, usually make the operation possible.

When the system under consideration is autonomous i.e. $\{f(r)\} \equiv 0$ the problem becomes that of reducing the vector $\{x(t)\}$ to the null vector. Equivalently the system must "hit" the origin of the state space. From equation (4-10) it can be seen that this situation will have been achieved when:

$$-\{x(0)\} = \int_0^{t}[H]^{-1}[B]\{u\}\, d\tau = \int_0^{t}[Y\ (\tau)]\{u(\tau)\}\, d\tau \qquad (4\text{-}11)$$

A number of things must now be proven. For example it must be shown that there exists a $t_o > 0$ for which the equation (4-11) is satisfied for any $u(t)$. Then it must be demonstrated that of all the different values of t_o that will satisfy this equation one of them, t^*, will be minimized by a suitable choice of $u(t)$. Finally the form of $u(t)$ must be determined.

The various points to be proven have been examined rigorously in the literature [20], (Bellman [69]), (Kurzweil [70]) and the proofs, which are reasonably lengthy and difficult, will not be reproduced here. The proofs indicate, however, that a vector $\{\eta\}$ must be found such that the "dot" product below is maximized.

$$\{\eta\} \cdot [Y]\{u\} \qquad (4\text{-}12)$$

This is the same as maximizing:

$$\{\eta\}'[Y]\{u\} \qquad (4\text{-}13)$$

It is shown in the literature referenced above that this maximum exists and will be achieved when:

$$\{u\} = \operatorname{sgn}\{\eta\}'[Y] \qquad (4\text{-}14)$$

$$\text{i.e. } |u_i| = 1$$

The procedure described above can be demonstrated by means of the following example.

Consider the case of a linear oscillator with the circuit configuration shown in Figure 4.10. This particular example is chosen as it is one discussed by Bushaw ([66], p. 42). His detailed discussion leads to the correct switching boundaries (ref [13], p.42) but with rather more effort than is required using the method described above.

The system equations in matrix form are:

$$\begin{Bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{Bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{Bmatrix} x_1 \\ x_2 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 1 \end{Bmatrix} u \qquad (4\text{-}15)$$

and it is known that for time optimum response the function u can take on the values of +1 or -1. In order to obtain the matrix solution the method of Lagrange can best be applied here. The solution to the matrix equation below must be found first:

$$[\dot{H}] = [A]\,[H] \quad \text{with} \quad [H(0)] = [I] \qquad (4\text{-}16)$$

which in this case becomes:

$$\begin{bmatrix} \dot{H}_{11} & \dot{H}_{12} \\ \dot{H}_{21} & \dot{H}_{22} \end{bmatrix} = \begin{bmatrix} H_{21} & H_{22} \\ -H_{11} & -H_{12} \end{bmatrix} \qquad (4\text{-}17)$$

This matrix equation yields four second order differential equations which can be solved with the initial conditions to give:

$$H_{11} = \cos t, \quad H_{12} = \sin t \quad H_{21} = -\sin t \quad \text{and} \quad H_{22} = \cos t$$

$$\text{or} \qquad [H] = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$$

therefore

$$[H]^{-1} = \begin{bmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{bmatrix} \quad \text{and} \quad [Y] = [H]^{-1}[B] = \begin{Bmatrix} -\sin t \\ \cos t \end{Bmatrix} \qquad (4\text{-}19)$$

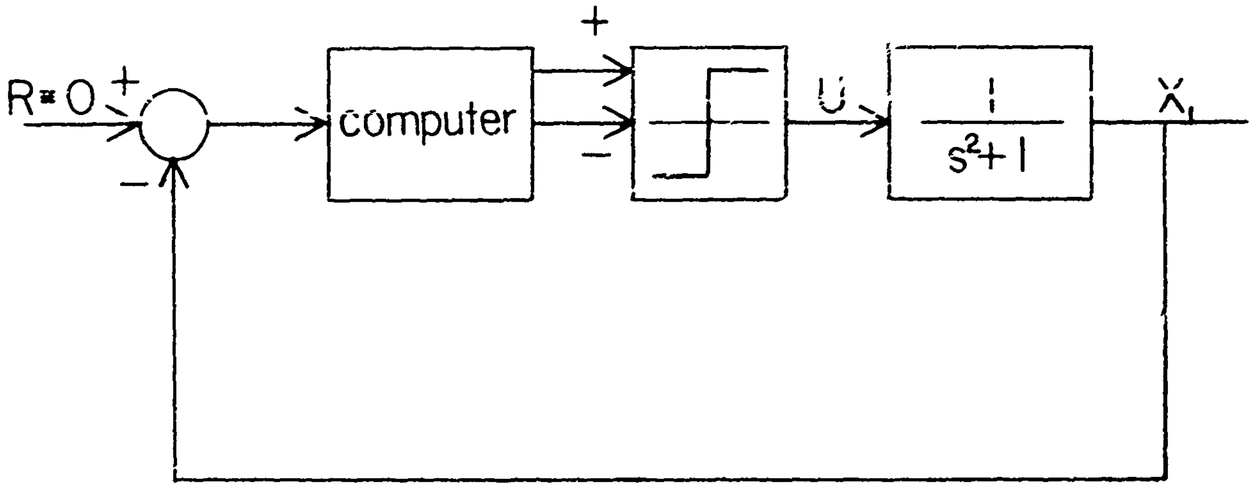

Figure 4.10

System with the Linear Portion a

Linear Oscillator

and $\{\eta\}'[Y] = -\eta_1 \sin t + \eta_2 \cos t$ (4-20)

which can also be written:

$\{\eta\}'[Y] = A\cos(t + d)$ A and d are functions of η_1, η_2 (4-21)

So it is seen that:

$u = +1$ when $A\cos(t + d) > 0$

and $u = -1$ when $A\cos(t + d) > 0$ (4-22)

and it is immediately apparent that the relay must switch every π seconds.

If the system is to reach the origin of the state space the final solution trajectory must go through the origin. There will be two such final trajectories, one for each of the two cases, u = +1 and u = -1. Furthermore, since these trajecteries are final trajectories, the system must "switch onto" them sooner or later and they are therefore part of the system switching boundary.

To find the final trajectories a technique suggested by La Salle [20] and others can be used. Let T = -t in the system equations, equation (4-15), and solve the equations with the initial conditions (0,0), i.e. let time run backwards away from the final point, the origin. Allowing time to run backwards for π seconds will generate that portion of the switching boundary through the origin. With this substitution equation (4-15) becomes in component form:

$$\frac{dx_1}{dT} = -x_2 \qquad (4\text{-}23)$$

$$\frac{dx_2}{dT} = x_1 - u$$

When u = +1 we get

$$x_1(T) = 1 - \cos T$$

$$x_2(T) = -\sin T \quad \text{or eliminating T} \quad (x_1 - 1)^2 + x_2^2 = 1 \qquad (4\text{-}24)$$

When $u = -1$ we get

$$x_1(T) = -1 + \cos T$$

$$x_2(T) = \sin T \qquad \text{or eliminating } T \quad (x_1 + 1)^2 + x_2^2 = 1 \qquad (4\text{-}25)$$

Considering the signs in the parametric equations and allowing T to increase to π seconds, two semi-circles result as shown in Figure 4-11. The direction of increasing t or decreasing T is toward the origin. Choosing an arbitrary final switching point on this portion of the boundary, say at (1, -1) for conveniency, the solution trajectory immediately prior to this final switching can be determined from the equations (4-23) with initial conditions (1, -1) and with $u = -1$. The result is:

$$x_1(T) = -1 + 2\cos T + \sin T$$

$$x_2(T) = 2\sin T - \cos T \text{ or eliminating } T \ (x_1 + 1)^2 + x_2^2 = 5 \qquad (4\text{-}26)$$

This trajectory is drawn in dotted lines on Figure 4.11. The parametric equations again indicate the direction of increasing t (decreasing T), thus showing that the portion of the state plane above the switching boundary found so far corresponds to $u = -1$, and the portion below to $u = +1$. Allowing time to run backwards along the dotted trajectory for π seconds determines one point on the switching boundary to the left of the existing portion. This point is labeled B in Figure 4.11. Constructing trajectories from different initial points on the known portions of the switching boundaries thus will yield additional portions of the boundary. The complete picture is clearly exactly that of Figure 4.8. It is to be observed that the components of the maximizing vector $\{\eta\}$ do not have to be found explicitly, but rather the possible behavior of Equation (4-14) is observed and the position of the boundaries is deduced from the zero crossings of $\{\eta\}'[Y]$.

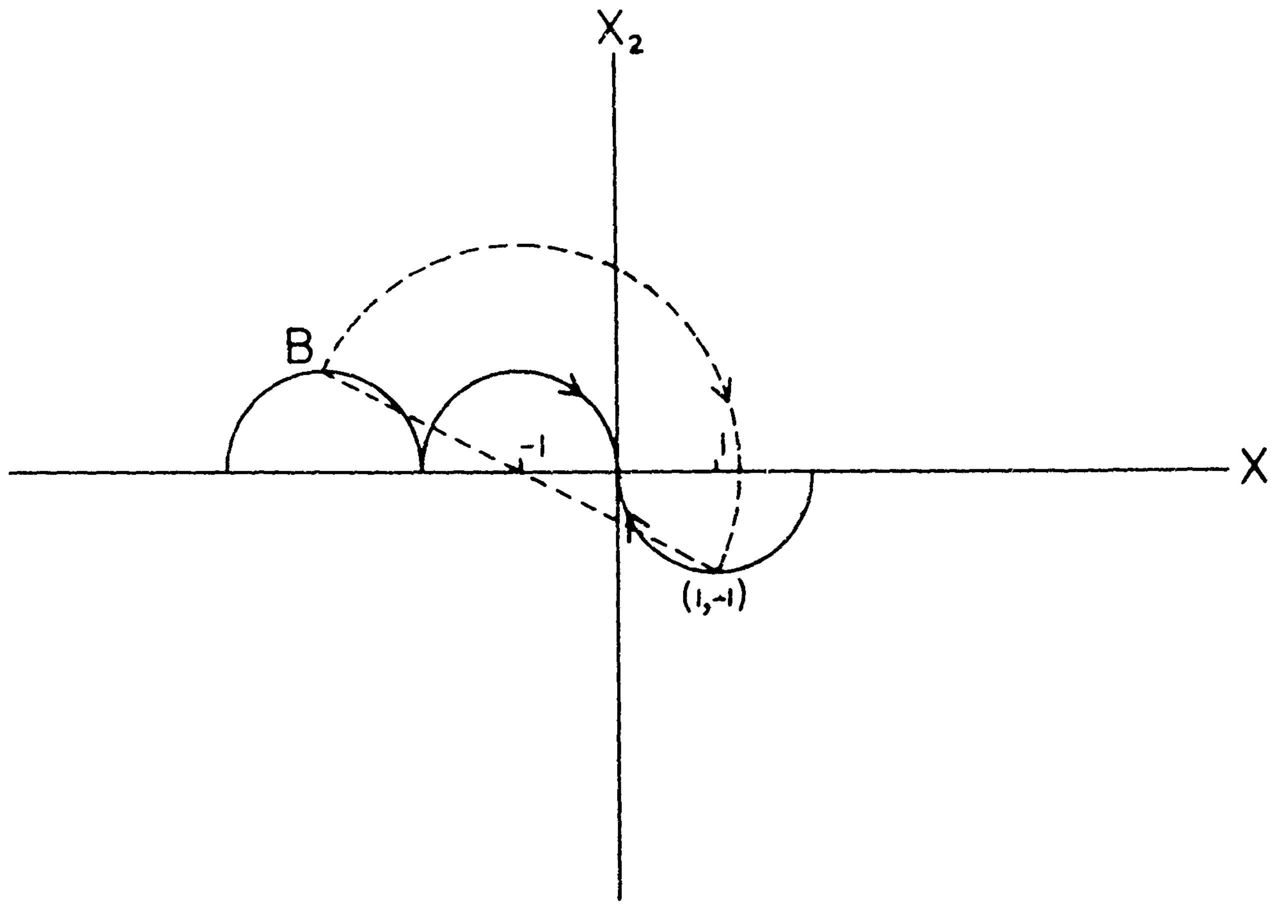

Figure 4.11

Construction of Switching Boundaries

for Linear Oscillator Circuit

The method outlined in this section is a method that will lead to switching boundaries which thus defines a time optimum policy from any admissable starting point in the phase space. The method does allow for solution of systems where the linear portion is described by a linear time varying equation. The solution is only possible then in the restricted cases when equation (4-6) can be solved.

The boundaries of the example system and of any more complicated system must be instrumented in the phase space. The position of the system with respect to these boundaries must be determined continuously in order that the relay can be switched to the correct polarity. There is also the additional problem that frequently not all the physical variables are available from which the state variables must be calculated. In this situation a prediction or estimation of the missing variables must be attempted (Kalman [71]).

The sequential procedure of boundary determination, boundary instrumentation, and the determination of system position via measurement or estimation is a complicated task even for simple systems. In addition, pre-determination of the optimum policy does not allow for unpredicted changes or disturbances. The concept of a multistage decision process (dynamic programming) (Bellman [72]) suggests that the necessary steps mentioned above should be undertaken repetitively by one computing element. The system optimum policy would then be considered as the solution of a synthesis problem that can be reviewed from moment to moment as the solution proceeds. These latter ideas are discussed in more detail in the next section.

4.3 The Synthesis Problem

A basic problem associated with switched systems is that of determining the switching boundaries. These are the hyper-surfaces located in the system state space where the system steering function must change sign in order to adhere to the optimum policy.

The technique outlined and referenced in the last section will solve the problem of finding the form of the steering function u as a function of time. In the case of autonomous systems and for time optimum response it is shown that:

$$u = \mathrm{sgn} \; \{\eta\}'[Y] \tag{4-27}$$

where the matrix $[Y]$ is time dependent. In order that the system shall satisfy this relation it is necessary, therefore, to determine the points in the state space where the sign of the function changes. To find the hyper-surfaces made up of all such points, the form of the steering function, already known as a function of time, must be determined as a function of the state variables. It is sometimes written that, the zero crossings of u, as a function of time, must be mapped into the state space. This is the desired situation, as it is assumed that the state variables can be constructed one way or another [71] and hence are available to feed the computing elements of Figures 4.6 and 4.10. The determination of the switching hyper-surfaces as functions of the state variables will be defined as the synthesis problem and this part of the theory of optimum systems is considered as a problem separate from the problem of determining the form of the function u as a function of time.

If the optimum policy can be determined with certainty ahead of time, as was the case with the simple example concluding the last section, the

computing element may construct and then monitor the system state variables. The computing element will decide where the system solution is in the state space with respect to the switching boundaries, and make a simple decision as to the optimum relay position as the solution proceeds.

Instrumentation in these cases is possible in a number of ways, e.g. two variables can be applied, one to each of the deflection plates of a calibrated oscilloscope. The tube face is then masked to match a switching boundary and monitored by a photo-electric cell (Hopkin [73]). Techniques suitable in these cases can be compared with the technique of pre-programming.

A more powerful approach to the problem exists, however, that can be used to take care of the situation when the optimum policy is known initially. The approach will also take care of situations when the optimum policy may well change as the solution proceeds. The approach is that known as dynamic programming [72].

The title dynamic programming is a phrase coined to describe the procedures associated with the solution of a multi-stage decision process. The procedure involves sampling the system position in the state space repetitively. At each sampling the computer makes a decision as to what the optimum policy is at that time. For example, there are two choices available in the case of a time optimum system with a single relay. The optimum policy decided at one sampling time is pursued until the process is repeated at the next sampling time.

Optimum systems defined in terms of their performance indicies that are to be treated in this fashion must possess the basic property of being Markovian, i.e., after any number of decisions, say k, the effect of the remaining

N-k stages of the decision process upon the total return must depend only upon the state of the system at the end of the k-th decision and the subsequent decisions ([54], p. 54). Systems that are Markovian in nature will then perform in a manner that is optimum overall, even when the decisions are made repetitively as the solution proceeds. The Markovian property is fortunately characteristic of most systems encountered.

It is to be observed that the past history of the system need not be considered in determining the future policy, and consequently such a procedure will allow for unpredicted disturbances, etc.

Instrumentation of the computer element is clearly no longer possible by simple means. In fact the suggested procedure has only become feasible with the advent of high speed digital computers which inevitably perform the computing task. There is still a finite computing time associated with the decision process, of course, and that places a limitation on the sampling frequency and in turn on the system performance.

An example of a method where the optimum policy is reviewed as the solution proceeds has been presented recently by Smith [74] and the technique will be summarized here.

Equation (4-10), the solution equation for the state variables, is reproduced here but with $f(r) = 0$:

$$\{x\ (t)\} = [H]\{x(0)\} + [H\ (t)]\int_0^t [H(\tau)]^{-1}[B(\tau\ )]\ \{u(\tau\ )\}\ d\tau \tag{4-28}$$

Allowing that in a time optimum problem the function u can only take on the values plus or minus one and that this function will change sign according to equation (4-27) at times t_1, $t_2 \ldots t_{n-1}$,T where $0 < t_1 \ldots < T$, equation (4-28) can be written:

$$[H(t)]^{-1}\{x(t)\} - \{x(0)\} = \pm\left[\int_0^{t_1}[H]^{-1}[B]\{u\}d\tau + \int_{t_1}^{t_2}\ldots\ d\tau + \ldots + \int_{t_{n-1}}^{T}\ldots d\tau\right] \tag{4-29}$$

If the set of n equations represented by the matrix notation are integrated, the result will be n simultaneous, algebraic equations with the times $t_1 \ldots\ldots T$ as unknowns. The solution of simultaneous equations on a digital computer is possible by means of well known techniques. Consequently two values of T, say T^+ and T^- corresponding to the choice of sign outside the bracket on the right hand side of Equation (4-29), can be calculated with knowledge of the initial conditions. The computer can now make a decision whether $T^+ > T^-$ or $T^- > T^+$, and can use the result of this decision to activate the relay in the appropriate direction. This procedure can be repeated continuously by sampling the system at intervals to obtain new initial conditions, i.e. the values of the state variables at the time of sampling. The sampling frequency can be made as small (or as large) as is desired, weighting the computing time against the system response time.

A number of other authors working with different forms of integral type performance indicies have sketched methods for the continuous solution of the synthesis problem. For example Lee [56] indicates possible methods for the continuous solution of minimum-energy, minimum-time and minimum-error systems.

The results that are given by the authors indicate that the desired response is possible with the techniques suggested. It is to be observed that the performance is obtained at some expense for the amplifier of the conventional system is replaced by a relay and a digitial computer!! This

fact is not, however, of direct concern if the objective in reviewing these methods is to determine the best possible system, for purposes of system evaluation. The physical construction of an optimum system is a separate problem which has been solved by one author, as described in this section.

4.4 Conclusions

The methods outlined in these sections depend to a large extent on the form of the system and on the form of the optimum response desired. The time optimum systems are those where the steering function is chosen to minimize the upper limit of the integral in equation (4-1). At the present state of the art much attention has been given to the time optimum problem. The solutions that do exist, however, are restricted almost entirely to systems of second order. The exception discussed in the last section used the full capacity of a large digital computer to handle a fourth order system. It is thus clear that for systems other than the simplest systems the existing methods are far from economic and cannot be considered as practical as yet.

The concept of using an optimum system as the upper bound on all systems that are to fulfill a given task is not restricted to the time optimum case. For example other optimum systems require that u be chosen with a fixed upper limit on the integral to minimize (or maximize) the integral itself. Alternative methods that may lead to the solution of these problems must be sought.

It has been shown (Rozonoer [75]) that all problems of this nature can be interpreted as a problem of maximizing (or minimizing) a coordinate in the state space, together with some constraints. For example if there is a system described by n equations as written in equation (4-4) and the optimum system is defined as a time optimum system with the additional constraint:

$$\int_0^T F(x,u,\tau)\,d\tau \leqslant A \qquad (4\text{-}30)$$

one can define two new variables, x_{n+1} and x_{n+2}. These variables would be defined:

$$x_{n+1} = t \qquad \text{i.e.} \qquad \dot{x}_{n+1} = 1$$

$$x_{n+2} = \int_0^t F(x,u,\tau)d\tau \qquad \text{i.e,} \quad \dot{x}_{n+2} = F(x,u,t) \qquad (4\text{-}31)$$

The system equations can now be re-written in the form of equation (4-4) but with n + 2 variables.

The requirement for the system to be optimum is now that the coordinate x_{n+1} is to be minimized with the constraint that the coordinate $x_{n+2}(T) \leqslant A$.

Systems of this type, where a coordinate or a linear combination of coordinates must take on an extreme value, can be optimized by the Principle of Pontryagin [76] (with a unified procedure). While the Maximum Principle of Pontryagin is in the hands of the applied mathematicians, at the present state of the art, it is to be hoped that, with attention from qualified engineers, it will be proven to be of great practical value in the design of optimum automatic control systems.

In conclusion, it is remarked that the methods for the determination and construction of the configuration necessary for an optimum system of any sort are as yet in their infancy for anything but the simplest systems. The methods reported in this chapter, however, indicate that in many cases unique optimum configurations exist in the mathematical sense of the word. The techniques associated with the method of dynamic programming and/or the Maximum Principle of Pontryagin seem to provide the paths along which practical solutions will be found. This is an area in which there is a

considerable interest from the engineering world at the present time. Further work will undobtedly prove to be of considerable value for the better evaluation and construction of automatic control systems.

CHAPTER V

THE RESPONSE OF FORCED NONLINEAR SYSTEMS

5.1 Introduction

While the stability and response of autonomous nonlinear systems is of considerable interest to the automatic control specialist and to a number of applied mathematicians throughout the world simply for the sake of the problem, the user of a control system could hardly be less interested. The unforced system is of no conceivable engineering use. For a system to be of engineering interest it must function properly over a specified range of inputs.

This report has considered the unforced system, not because of any misapprehension that such systems are of any practical use, but rather because such studies should lead to a better understanding of the forced system. It is hoped that the techniques used for forced systems may be extended to the analysis of unforced systems. Such extensions are not easy of course. Let us cite two examples of different kinds of problems that arise. First, it has been shown above that the phase plane analysis may be extended to an approximate analysis of a forced time-invariant second order system. In this example the only difficulty arises from the approximation required of the input and the increased labor required to obtain the result. A second example however, will bring out a more basic difficulty. In the discussion above on the DIDF it has been shown that the use of the conventional DF for obtaining closed loop frequency response is fraught with danger. Examples can be given in which the conventional DF yields an apparently perfectly satisfactory closed loop frequency re-

sponse for a system that is actually unstable. This is the second class of problem. Not merely is the method of analysis slightly more cumbersome in forced systems, but rather it yields an incorrect answer.

Turn now to the problem of direct concern; the forced nonlinear system. This research is concerned with the development of specifications for control systems. Clearly the AF is not interested in the detailed nonlinearities involved in a particular system. The AF is interested in performance, and the particular techniques involved in meeting the specifications must remain in the design province. Thus it would seem improper for this group to become involved in compiling a list of all possible nonlinearities and specifications for them unless this were the sole method open. The approach that appears of most validity would seem to be the establishment of the optimum performance within the given set of constraints for a given situation. This theoretical optimum may then be used as an Index of Performance with which to judge the performance of competing physical systems.

It must be asked why this approach is suggested when it obviously failed with linear systems. Considerable effort was extended in attempts to obtain general Indices of Performance for linear systems. It will be recalled that such IP's were sought in terms of system parameters. In this new effort the IP will be formulated in terms of the constraints on the system. Such constraints might be maximum force or torque, maximum power or a finite stored energy with which to accomplish the task. In effect these are the basic and fundamental nonlinearities of the system, and meaningful maximum performance values can be formulated within them. Such constraints do not exist in linear systems, of course, and so it was impossible to formulate such a policy concerning them in the linear portion of this study. Since

physical constraints exist universally in practice, this new approach should be of wide applicability.

To establish the optimum system within a given set of constraints it is necessary to define the sense in which the word optimum is used. Most of the work done to date on the optimum problem concerns the time optimum case. The desired task is to reduce the error to zero in minimum time. This is a resonable approach, although perhaps not the simplest, and is discussed in Section 5.3. In Section 5.2 the conventional phase plane analysis is extended to the approximate calculation of the response to an arbitrary input. Of course the practical limitation of the phase plane to second order systems severely limits the usefulness of this approach. It was not thought necessary to include a separate section in this chapter for the DIDF. This technique is definitely of use in finding the response of a nonlinear system to a sinusoidal input, but the method has already been discussed in the chapter on the stability of forced systems.

In Section 5.4 the response of a nonlinear system to random noise is discussed. At first glance this might seem to be a digression from the main stream of this survey. It must be pointed out that this is not so.

Unfortunately the carry-over of linear system concepts and points of view in the thinking of many investigators when they turn to nonlinear systems seems doomed to failure. This persistance would appear to be the only logical answer, for example, in the continued and unwarranted attention paid to the response of a nonlinear system to a sinusoidal driving function. The sine wave no longer reigns supreme in nonlinear system analysis; it is reduced to just another special input that provides no special insight to the overall behavior of the nonlinear system.

Wiener has proposed that the basic and most general input for a nonlinear system is the random signal provided by Brownian motion. He points out that from the response of a nonlinear system to Brownian motion, it is possible to predict its response to any input; much as it was possible with the response to sine waves in linear systems. Thus Brownian motion input occupies the same position with respect to nonlinear systems as does the sinusoidal input to linear systems. This concept appears very powerful, although it is not yet in an operational condition, and it is discussed in Section 5.4.

5.2 The Phase Plane for Forced Systems

The phase plane can be used to calculate and display the response of forced second order systems (Gibson [77]). The approach is, of course, not as simple as with autonomous systems but is still quite practical. The input may be arbitrary and the approach makes use of all of the convenience of construction methods, such as the delta method and isocline method, that are available for autonomous systems.

The method is based on representing the actual input by a train of equivalent impulse functions. It is also necessary to represent the system nonlinearity in an equivalent piecewise linear fashion. Each impulse is considered as an initial condition for the interval between it and the next pulse, superimposed on the final condition of the system at the end of the preceeding interval. The central convenience of this approach lies in the fact, as pointed out by Trimmer [78] , that in a second order system an impulse appears simply as a sudden increment of velocity in the system. Thus no change need be made in the isoclines or delta function used in the construction of the phase portrait of the autonomous system. The method

appears to be of promise as an approximation technique for forced second-order systems but requires further experience before it can be recommended as a method for obtaining proof of compliance to performance specifications. Gibson [77] works several examples, but his presentation is not complete or exhaustive.

5.3 Time Optimum (Switched) Nonautonomous Systems

Although most studies of the optimum control problem restrict themselves to the autonomous system, a few workers have considered the more difficult problem of time-optimum nonautonomous systems. Amoung these are Krasovskii [79], Kalman and Koepcke [80], LaSalle [20] and Fuller [81].

LaSalle has developed a number of basic theorems for forced, time varying parameter, time optimum control systems, but he does not discuss the engineering implications of his work. Kalman and Koepcke point out certain of the implications of such systems, and Fuller further clarifies the situation.

It is no longer sufficient in general to use only the state variables of the autonomous system for the description of the forced system. The state variables to be fed into the logic element or controller must now consist of a state description of the plant plus a state description of the input

$$\left\{\overset{\circ}{x}(t)\right\} = \left[A(t)\right]\left\{x(t)\right\} + \left[B(t)\right]\left\{u(t)\right\} + \left\{f(t)\right\} \qquad (5\text{-}1)$$

It must be pointed out that this formulation of the equations of a non-linear closed loop system <u>is not complete</u>. A further nonlinear relation will be required to relate $\left\{x(t)\right\}$ and $\left\{u(t)\right\}$. For example in the system in Figure 3.1, it will be found convenient to identify the state variables with error, the steering function with m, and the forcing function with the input. Thus in matrix form

$$\left\{\dot{e}(t)\right\} = \left[A(t)\right]\left\{e(t)\right\} + \left[B(t)\right]\left\{m(t)\right\} + \left\{h(t)\right\} \tag{5-2}$$

where $f(t)$ of equation 5-1 is now say $g(r) = h(t)$

Now $m = f(e)$ say. It will be impossible to substitute this relation into the matrix formulation and manipulate because matrix algebra is a linear algebra. Thus the matrix relation is an open loop description of a portion of the system. When the nonlinear relation is considered after the matrix manipulations are complete, the most difficult problem still remains. Most authors ignore this problem completely although a few of the bravest acknowledge its existence. None solve it.

In order to define the optimum trajectory and the optimum switching boundaries in the methods discussed in the literature the input must be defined in the form of a differentiable function, e.g. a polynomial. This is a restriction. Note that under these conditions the optimum switching boundaries are directly dependent upon the description of the input signal. This means that the complexity of the system will rise rapidly if this approach is followed with the result that all insight will be lost. Secondly, since the optimum design is uniquely tied to the defined input state coordinates, even if it could be instrumented, the system would be useful for only the one input for which it was designed. Of course this situation is impossible. Fortunately, however, several alternative schemes present themselves.

Kalman and Koepcke [80] suggest that the actual input be approximated with a curve of given degree over given segments of time. Then the coefficients of the approximating polynomial need be found only once per time segment. These numbers could be entered into the optimum switching curve calculations, thus, the switching boundaries could be recalculated once each

segment. The concept of switching boundaries in the state space for a nonautonomous system appears in reality, however, to be a rather cumbersome one and perhaps deserves to be abandoned for a more sophisticated approach. Bellman suggests that the concept of a multistage decision policy or dynamic programming is directly applicable here [54].

The philosophy of approach is as follows. A general performance index or payoff function is defined in terms of the system state variables and the state description of the input.

$$IP = \int_0^\infty Q \quad (x_1, x_2, x_3, \ldots x_n, r_1, r_2, \ldots, r_n) dt \tag{5-3}$$

At given instances of time the question is asked, which state of the relay will optimize the IP at this time? The question is asked sequentially in time and the relay adjusted accordingly as the solution proceeds. So long as the system, described in terms of the state variables, possesses the Markovian Property ([54], p. 54) (that is, we do not have time delays nor a "delay differential equation" to describe the system) Bellman has shown that a sequence of such optimum decisions is optimum overall. Naturally the engineering implementation of such a scheme is considerably more complex than this naive description might indicate, and, in fact, engineering feasibility studies of it are only now being initiated in various engineering research laboratories. Without doubt, this is a most challanging area for future research and ties in directly with the already wide spread activity on the concept of adaptive control.

5.4 Response of Nonlinear Systems to Random Inputs

The response of a certain class of nonlinear systems to random inputs

has been investigated by some researchers. Considerable effort has been expended in investigating the theory of nonlinear filters. Wiener [82] considers the response of a particular type of nonlinear filter to Brownian motion input. The type of filter to which Wiener's method is applicable is the one in which it is possible to separate the filter into two distinct blocks in tandem, one block being linear and the other nonlinear of the functional, instantaneous (non memory) type. Wiener shows that the knowledge of the Brownian motion response of such a system is sufficient for the determination of its response to any other input. The application of this theory to the investigation of the response of a nonlinear feedback system is not known at the moment. Bose [83] also considers the problem of determining the response to random inputs of a filter similar to the one treated by Wiener.

The determination of some statistical characteristics of the output of certain nonlinear nonclosed loop systems with random inputs is considered by Rice [84], Laning & Battin [85] and Sridhar [53].

The problem of analysis and synthesis of nonlinear feedback systems subject to stochastic inputs is certainly a wide open area for research and warrants considerable effort. This is evident when one considers the work of Wiener, since this provides a fresh viewpoint to the whole problem of analysis and synthesis of nonlinear systems. At the present Wiener's work is not applicable to any but the most trivial practical cases.

BIBLIOGRAPHY

1. W. J. Cunningham, Introduction to Nonlinear Analysis, McGraw-Hill Book Co., New York; 1958.

2. J. J. Stoker, Nonlinear Vibrations, Interscience Pub. Inc., New York; 1950.

3. N. Minorsky, Introduction to Nonlinear Mechanics, Edward Bros. Pub., Ann Arbor, Mich.; 1957.

4. A. A. Andronov, C. E. Chaikin, Theory of Oscillations, Princeton U. Press; 1949.

5. S. Lefschetz, Differential Equations: Geometric Theory, Interscience Pub., New York; 1957.

6. R. E. Bellman, Stability Theory of Differential Equations, McGraw-Hill Book Co., New York; 1953.

7. La Salle, J. P., "Some Extensions of Liapunov's Second Method," RIAS Technical Report No. 60-5; 1960.

8. N. Krylov and N. Bogoliuboff, "Introduction to Nonlinear Mechanics," Princeton University Press, Princeton, N. J.; 1943.

9. L. C. Goldfarb, "On Soms Nonlinear Phenomena in Regulator Systems," Automatikai i Telemekanika, vol. 8, No 15, pp. 349-383; 1947.

10. R. J. Kochenburger, "A Frequency Response Method for Analyzing and Synthesizing Contactor Servomechanisms," Trans. AIEE, vol. 69, part 1, pp. 270-284; 1950.

11. A. Tustin, "The Effects of Backlash and of Speed-development Friction on the Stability of Closed-Cycle Control Systems," Journal of IEE, vol. 94, pt. 21, No. 1; 1947.

12. W. Oppelt, "Focus Curve Method for Regulators with Friction," J. Inst. E. E., vol. 94, Part 2A, Nos. 1 and 2; May 1947.

13. E. C. Johnson, "Sinusoidol Analysis of Feedback-Control Systems Containing Nonlinear Elements," Trans. AIEE, vol. 71, part 2, A & I, pp. 169-181; 1952.

14. J. C. Gille, M. J. Pelegrin and P. Decaulne, Feedback Control Systems, McGraw-Hill Book Co., Inc., N. Y.; 1959.

15. E. Levinson, "Some Saturation Phenomena in Servomechanisms with Emphasis on the Tachometer Stabilized System," Trans. AIEE, vol. 72, part 2, A and I, pp. 1-9, March, 1953.

16. C. Hill, "A Closed Loop Frequency Response Method for Non-Singlevalued Nonlinearities," to be published.

17. J. Z. Tsypkin, Teorija Relejmykh Sistem Automaticheskovo Regulirovanija, Gostekhizdat, Moscow; 1956.

18. Y. Ku, Analysis and Control of Nonlinear Systems, Roland Press Co., New York; 1958.

19. L. A. Pipes, Applied Mathematics for Engineers and Physicists, McGraw-Hill Book Co., New York; 1958.

20. J. P. La Salle, "The Time Optimal Control Problem," Contributions to the Theory of Nonlinear Oscillations, vol. 5, Princeton University Press; 1960.

21. R. Bellman, Introduction to Matrix Analysis, McGraw-Hill Book Co., New York; 1960.

22. R. E. Kalman, "Analysis and Design Principles of Second and Higher Order Saturating Servomechanisms," AIEE Trans., part 2; Nov. 1955.

23. J. E. Gibson, et al, "Specification and Data Presentation in Linear Control Systems," vol II, Final Report on Project AF 29(600)-1933, Purdue University; May 1961.

24. W. Kaplan, Ordinary Differential Equations, The Adison-Wesley Publishing Co., Reading, Mass.

25. I. G. Malkin, Theory of Stability of Motion (translation from Russian) Atomic Energy Commission Translation No. 3352; 1959.

26. V. I. Zubov, Mathematical Methods of Control System Analysis, (in Russian) Sudpromizdat; 1959.

27. W. Hahn, "Theory and Application of the Direct Method of Liapunov" (in German) Ergebnisse der Mathematik, vol. 22, Springer Verlag, Berlin; 1959.

28. R. E. Kalman, J. E. Bertram, "Control System Analysis and Design via the "Second Method" of Liapunov," ASME Paper No. 59-NAC-2; 1959.

29. L. F. Kazda, Editor, "Work Session in Liapunov's Second Method," University of Michigan, Ann Arbor; 1960.

30. J. E. Gibson, et al, "Stability of Nonlinear Control Systems by The Second Method of Liapunov," vol III, Final Report on Project AF 29(600)-1933, Purdue University; May 1961.

31. A. I. Lur's, Some Nonlinear Problems in the Theory of Automatic Control, (translation from Russian), Her Majesty's Stationery Office; 1957.

32. A. M. Letov, Stability in Nonlinear Control Systems, (translated by J. G. Adashko), Princeton University Press; 1961.

33. Yu. I. Alimov, "Liapunov Functions For Relay Control Systems," Automatics and Remote Control, vol. 21, No. 6, pp. 720-728; 1960.

34. B. S. Razumikin, "The Application of Liapunov's Method to Problems in the Stability of Systems with Delay". Automatics and Remote Control, vol. 21, No. 6, pp. 740-748; 1960.

35. Z. V. Rekasius and G. P. Szego, "Stability Analysis of Systems with Limit Cycles," (to be published).

36. H. A. Antosiewicz, "A Survey of Liapunov's Second Method," Contributions to the Theory of Nonlinear Oscillations, vol IV, pp. 141-166; 1958.

37. G. P. Szego, "A Stability Investigation of Time-Varying Linear Systems," AIEE Conference paper number CP-61-741; 1961.

38. J. L. Massera, "Contributions to Stability Theory," Annals Math. vol 66, pp. 182-206; 1956.

39. R. Oldenburger, "Signal Stabilization of a Control System," Trans. ASME, vol 79, no 8, pp. 1869-1872; 1957.

40. R. Oldenburger and C. C. Liu, "Signal Stabilization of a Control System," Trans. AIEE, vol 78, part 2, pp. 96-100; 1959.

41. N. Minorsky, "On Asynchronous Action," Journal of Franklin Institute, vol 259, no 3, pp 209-219; 1955.

42. R. Oldenburger and T. Nakada, "Signal Stabilization of a Control System," presented at IRE National Convention at Dallas, Texas; Nov. 1959.

43. R. Oldenburger and R. Sridhar, "Signal Stabilization of A control System with Random Inputs," to be presented at Joint Automatic Control Conference, Boulder, Col.; June 1961.

44. R. Sridhar and R. Oldenburger, "Stability of a Nonlinear Feedback System in the Presence of Gaussian Noise," to be presented at the Joint Automatic Control Conference, Boulder, Col.; June 1961.

45. R. Oldenburger and R. C. Boyer, "Sinusoidal Signal Stabilization," to be presented.

46. J. E. Gibson and R. Sridhar, "A New Dual Input Describing Function," to be presented.

47. E. P. Popov, "Dynamik Automatisher Regelsysteme," (Dynamic Automatic Control Systems), section 60, Akademie-Verlag, Berlin; 1958.

48. W. A. Stein and G. J. Thaler, "Obtaining the Frequency Response Characteristic of a Nonlinear Servomechanism for an Amplitude and Frequency-Sensitive Describing Function," Trans. AIEE, vol 77, part 2, pp. 91-96; 1958.

49. K. Ogata, "An Analytic Method for finding the Closed Loop Frequency Response of Nonlinear Feedback Control Systems," Trans. AIEE, vol 76, part 2, pp. 277-285; 1957.

50. L. T. Prince Jr., "A Generalized method for determining the Closed Loop Frequency Response of Nonlinear Systems," Trans. AIEE, vol 73, part 2, pp. 217-224; 1954.

51. J. C. West, J. L. Douce, and R. K. Livesey, "The Dual-Input Describing Function and Its Use in the Analysis of Nonlinear Feedback Systems," Proc. IRE, vol 103, pt. B, pp. 463-474; 1956.

52. R. C. Booton, Jr., "The Analysis of Nonlinear Control Systems with Random Inputs," Proc. Sym. Nonlinear Circuit Analysis, vol II, Brooklyn Polytechnic Institute, pp. 369-391; 1953.

53. R. Sridhar, "Signal Stabilization of a Control System with Random Inputs," Ph.D. thesis, Purdue University; January 1960.

54. R. Bellman, Adaptive Control Processes, Princeton University Press, Princeton, N. J.; 1961.

55. C. W. Merriam, "A Class of Optimum Control Systems," Journal of the Franklin Institute, vol 267; April 1959.

56. E. B. Lee, "Design of Optimum Multivariable Control Systems," Journal of the ASME, Series D; March 1961.

57. A. R. Forsyth, Calculus of Variations, Dover Publications Inc., New York 14, New York.

58. I. Bogner and L. Kazda, "An Investigation of the Switching Criteria for Higher Order Contactor Servomechanisms," Trans. AIEE, part 2; July 1954.

59. R. Oldenburger, "Optimum Nonlinear Control," Trans. ASME, vol 79; April 1957.

60. R. Neiswander and R. McNeal, "Optimization of Nonlinear Control Systems by means of Nonlinear Feedbacks," Trans. AIEE, part 2; September 1953.

61. D. A. Kahn, "An Analysis of Relay Control Systems," Trans. AIEE, vol 68; 1949.

62. K. Izawa and L. E. Weaver, "Relay type Feedback Control Systems with Deadtime and Sampling," Trans. AIEE, part 2; May 1959.

63. D. McDonald, "Multiple Mode Operation of Servomechanisms," Review of Scientific Instruments, vol. 23; January 1952.

64. R. Buland and N. Furomoto, "Dual Mode Relay Servomechanisms," Trans. AIEE, vol. 78; June 1959.

65. I. Flugge-Lotz, Discontinuous Automatic Control, Princeton University Press, Princeton, N. J.; 1953.

66. D. W. Bushaw, "Differential Equations with a Discontinuous Forcing Term," Ph.D. Dissertation, Department of Mathematics, Princeton University; 1952.

67. H. S. Tsein, Engineering Cybernetics, McGraw-Hill Book Co. New York; 1954.

68. S. S. L. Chang, "Optimum Switching Criteria for Higher Order Contactor Servo with Interrupted Circuits," Trans. AIEE, vol. 74, part 2; November 1955.

69. R. Bellman, I. Glicksberg and O. Gross, "On the Bang-bang Control Problem," Quarterly of Applied MatHematics, vol. 14; April 1956.

70. F. Kurzweil, "The Analysis and Synthesis of Nonlinear, Continuous, and Sampled Data Systems Involving Saturation," Stanford Electronics Laboratory, Stanford California, Tech. Report No. 2101-1; November 1959.

71. R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," Journal of the ASME, Series D; March 1960.

72. R. Bellman, Dynamic Programming, Princeton University Press, Princeton, N. J.; 1957.

73. A. M. Hopkin, "A Phase Plane Approach to the Compensation of Saturating Servomechanisms," Trans. AIEE, vol. 70; 1951.

74. F. B. Smith, "Time Optimal Control of Higher Order Systems," IRE, PGAC, vol. AC-6; February 1961.

75. L. T. Rozonoer, "L. S. Pontriagin's Maximum Principle in the Theory of Optimum Systems," Avtomat. i Telemekh. vol. 20; October 1959.

76. L. S. Pontriagin, "Optimum Control Processes," Uspekhi Matem. Nauk. vol. 13, Issue 1; 1959.

77. J. E. Gibson, Nonlinear Automatic Control, (in press) McGraw-Hill Book Co., New York.

78. J. R. Trimmer, Response of Physical Systems, John Wiley and Sons, Inc., New York; 1950.

79. N. N. Krasovskii, "Concerning the Theory of Optimal Control," Automation and Remote Control, vol. 18, pp. 1005-1016; 1957.

80. R. E. Kalman and R. W. Koepeke, "Optimum Synthesis of Linear Sampling Control Systems Using Generalized Performance Indices," Trans. ASME, vol. 80, pp. 1820-1826; 1958.

81. A. T. Fuller, "Optimization of Non-Linear Control Systems with Transient Inputs," Journal of Electronics and Control, vol. 8, No. 6, pp. 465-479; 1960.

82. N. Wiener, "Nonlinear Problems in Random Theory," Published Jointly by The Technology Press and John Wiley and Sons, Inc. 1958. (In particular refer to pages 88-100).

83. Bose, "Nonlinear System Characterization and Optimization," Trans. IRE, vol. CT-6, pp. 30-40; May 1959.

84. C. O. Rice, "Mathematical Analysis of Random Noise," Bell System Technical Journal, vol. 23, pp. 282-332; 1944, and vol. 24, pp. 46-156; 1945. Also reprinted in Wax, "Selected Papers in Noise and Stochastic Processes," Dover, New York; 1954.

85. J. R. Laning and R. H. Battin, "Random Processes in Automatic Control," McGraw-Hill Book Co., New York; 1956.

RECENT RESEARCH PUBLICATIONS
SCHOOL OF ELECTRICAL ENGINEERING
PURDUE UNIVERSITY

60-1 SURVEY OF THE PHILOSOPHY AND STATE OF THE ART OF ADAPTIVE SYSTEMS
G. R. Cooper, J. E. Gibson, et al., Technical Report No.1, Contract AF 33(616)-6890, July, 1960

60-2 THE SCATTERING OF A PLANE ELECTROMAGNETIC WAVE BY A FINITE CONE
C. C. Rogers and F. V. Schultz, Contract AF 19(604)-4051, August, 1960

60-3 A STUDY OF DIGITAL ADAPTIVE CONTROL SYSTEMS
J. T. Tou, P. D. Joseph, and J. B. Lewis, Contract Nonr 1100(18) NR 049-150, September, 1960

60-4 SPECIFICATION AND DATA PRESENTATION IN LINEAR CONTROL SYSTEMS
J. E. Gibson, E. S. McVey, C. D. Leedham, Z. V. Rekasius and R. Sridhar, Final Report, Vol. I, Contract AF 29(600)-1933, October, 1960

61-1 MICROWAVE ENERGY CONVERSION
E. M. Sabbagh, Part 1, Contract AF 33(616)-7355, WADD Technical Report 61-48, January, 1961

61-2 A CLASS OF ADAPTIVE CONTROLS
J. E. Gibson, Principal Investigator and J. S. Meditch, Final Report, Vol. I, Contract AF 33(616)-6890, February 1, 1961

61-3 LIMITS ON THE IDENTIFICATION TIME FOR LINEAR SYSTEMS
G. R. Cooper, Principal Investigator and J. C. Lindenlaub, Final Report, Vol. II, Contract AF 33(616)-6890, February 1, 1961

61-4 SPECIFICATIONS AND DATA PRESENTATION IN LINEAR CONTROL SYSTEMS-PART TWO
J. E. Gibson, E. S. McVey, C. D. Leedham, Z. V. Rekasius, D. G. Schultz, and R. Sridhar, Final Report, Vol. II, Contract AF 29(600)-1933, May, 1961

61-5 STABILITY OF NONLINEAR CONTROL SYSTEMS BY THE SECOND METHOD OF LIAPUNOV
J. E. Gibson, E. S. McVey, C. D. Leedham, Z. V. Rekasius, D. G. Schultz, and R. Sridhar, Final Report, Vol. III, Contract AF 29(600)-1933, May, 1961

61-6 PRESENT STATE OF THE ART IN THE SPECIFICATION OF NONLINEAR CONTROL SYSTEMS
J. E. Gibson, E. S. McVey, C. D. Leedham, Z. V. Rekasius, D. G. Schultz, and R. Sridhar, Interim Report 2, Contract AF 29(600)-1933, May, 1961